

Acquisition of Language Particular Properties under Impoverished Input

Bosook Kang, Ph.D.

University of Connecticut, 2005

This thesis investigates the acquisition of two language particular properties in Korean: Multiple Nominative Case marking (MNC) and scrambling. It is often assumed in language acquisition research that language input available to the child is fully representative of the adult grammar, providing samples of all the relevant properties possible in the adult grammar. That seems to be true for some well-known properties such as Wh-movement or Pro-drop. This thesis reports cases of other types which challenge such an assumption. We observe from the study of Korean data that the positive evidence for MNC and scrambling is drastically rare, raising a learnability question: how do children acquire those Korean language particular properties under the impoverished input?

By investigating the learnability question, we aim to answer the following related questions: i) To what extent does input frequency play a role in language acquisition? ii) Is the Principles and Parameters learning theory inevitable? If so, what language specific knowledge is required to ensure the acquisition of MNC and scrambling? iii) What kind of experience triggers the learning? We examine both child speech and parental input, and also consider syntactic theories of MNC and scrambling. The hypotheses gathered from production data and learnability implications of the syntactic theories of MNC and scrambling are finally tested in experiments.

The study reports several important findings. For both MNC and scrambling, we observe a very close resemblance between the child speech and the input patterns, giving the appearance that the pattern of language acquisition may be explained by an input-matching theory. However, we argue that such an impression is only apparent. My experimental results show that the child must make use of language specific hidden knowledge, providing support for the Principles and Parameters approach to language acquisition. The acquisitional findings also indicate that the triggering data may not necessarily be the sentences which directly exemplify a given property, arguing against input-driven learning approach. Further, we show that the areas of remarkable success of the child in matching input frequency are aspects of performance, not ones in which the child has to determine whether a given property is grammatical or not.

Bosook Kang, University of Connecticut, 2005

The Acquisition of Language Particular Properties
under Impoverished Input

Bosook Kang

B.A. Sookmyung Women's University, 1992

M.A. University of North Texas, 1999

M.A. University of Connecticut, 2003

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2005

UMI Number: 3187733

INFORMATION TO USERS

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 bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send 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.

UMI[®]

UMI Microform 3187733

Copyright 2006 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

APPROVAL PAGE

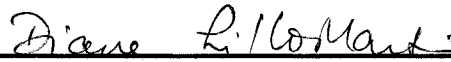
Doctor of Philosophy Dissertation

The Acquisition of Language Particular Properties under Impoverished Input

Presented by

Bosook Kang, B.A., M.A.

Co-major Advisor



Diane Lillo-Martin

Co-major Advisor



William Snyder

Associate Advisor



Željko Bošković

University of Connecticut

2005

Acknowledgments

To me, the process of writing a dissertation was a time of learning as much as taking courses or even more. The thesis has undergone several considerable transitions before it has the shape as it appears now. To be frank, I did not know where each step was leading. Fortunately, however, I found myself getting more interested in the questions and implications that emerged in the process. I would like to express my heartfelt thanks to those who made the journey endurable, worthwhile, and also enjoyable.

First, I would like to thank my three committee members: Diane-Lillo-Martin, William Snyder and Željko Bošković. Diane has been my pillar in the whole process. So many times, I would go to her office with half-cooked ideas or even just primitive intuitions on some issues. She always patiently listened to my unorganized mumblings, and shared with me her insight on those issues. The major questions that I addressed in the thesis came out of my discussions with her. Each meeting with her provided me with invaluable ideas with which to proceed to the next step of writing. She was the first person to read all the drafts, and gave me comments virtually on every point of the arguments. In addition to the academic advising, she was an amazing mentor. She never failed to make me feel encouraged whenever I left her office after the meeting. How grateful I felt every time! I received great help and support from William, too. His broad knowledge on language as well as on various other subjects always impressed me, let alone his expertise on statistics. Without his constructive comments and insightful suggestions, many parts of this thesis would have had serious errors. I am greatly indebted to Željko for his invaluable help with my understanding on syntactic issues. His lectures on the theory of scrambling kindled my first interest in scrambling and the acquisitional questions. I also deeply appreciate for his keen comments and his patience in dealing with some of my repeated mistakes in my draft.

Thanks also go to Judy Markus for her secretarial work for the department as well as for her genuine care about students.

I owe many thanks to other faculty members and students in the department. Yale Sharvit was always there with kind smiles, to help me with my questions on semantic issues, particularly in my second general paper on resumptive pronouns in child grammar. I was lucky to have an opportunity to work as a Ling 101 TA under Jonathan Bobaljik and Susi Wurmbrand, when they recently joined the department. Everything I learned about job application I owe to Susi's class on professional method. I enjoyed my last Ling 101 TA job under Harri van der Hulst in my last semester. Thanks are also due to Eva Bar-Shalom, who always gave encouragement and showed her interest in my research. Mona Anderson certainly made a unique mark in my memory for having a genuine interest in language.

I would also like to thank both students, who graduated before me, and those who are still in the program, for their contribution to creating a friendly academic environment: just to mention some of them, Kazumi Matsuoka, Debbie Chen, Kazuko Hiramatsu, Fumikazu Niinuma, Miguel Rodriguez-Mondonedo, Iseda Takako, Elaine Grolla, Simona Herdan, Jill Magee, Sara Felber, Inkie Chung, Bumsik Park, and Sei-Rang Oh, Claus Abel, Koji Sugisaki, Sandy Wood, Nilufer Sener, Karen O'brien and many more. Particularly, I thank Fumi and Miguel for their friendship and help with various things from discussing linguistic issues to technical matters. Kazumi Matsuoka was already a graduate when I entered the program. But she generously offered her time to discuss my research whenever she visited the campus. I would also like to thank a visiting scholar, Prof. Jeong-shik Lee's family for their warm hearted hospitality.

My thanks go outside the linguistics circle. The friends and volunteers in the International Christian Fellowship have been virtually my family for the last couple of years: Rev. Richard and his wife Carol, Steve and his wife Debbie, Frank and his wife Joan, Jon and his wife Ruth, Creig and his wife Nancy, Fernadel and his wife Maricel, Paul and his wife Mercy, Chichi, Chingching, Gloria, Sylvie, Akini, Marion, Inkyoo, Jinhee, Aaron, Onyi, and many more. Without their

prayers, love, care, guidance, and spiritual support, I would not have come this far. I would not exchange the joy and the fun that I had in the fellowship on every single Friday evening for anything. Especially, I give my heartfelt thanks to Billy and Mary, who have become my American family. They have made my otherwise solitary graduate life of living in the dorm enjoyable, making so many sweet memories for me, by teaching me how to play guitar and even by driving hundreds of miles to a theater in Pennsylvania. I would also like to thank my friends in Korea, Youngjin, Mijin, and Misoon for their long term friendship across the ocean. Last, but certainly not least, I am grateful to my parents, brothers and a sister for letting me study away from home for many years.

Above all, I give my utmost thanks to God for orchestrating all this in His mighty ways and for His grace to reveal to me who He is through all this time.

Table of Contents

Chapter 1: Theory of Language Acquisition

1.0. Introduction	1
1.1. Is frequency a determining factor in language development?	7
1.2. Could general inductive reasoning reconcile the mismatch?	12
1.3. Is it generalization at all?	16
1.4. Children's errors	19
1.5. UG-based approach	28
1.6. Outline of the dissertation	30

Chapter 2: Acquisition of Multiple Nominative Case Marking

2.0. Introduction	35
2.1. Case theory	37
2.2. Multiple Nom case in Korean	40
2.3. Two competing views on multiple Nom case	43
2.3.1. Default Nom approach... ..	43
2.3.2. Multiple feature checking approach	46
2.3.3. Implications for learnability	49
2.4. Acquisition of single Nom Case in Korean	51
2.5. Developmental contour of Nom case.....	56
2.6. Acquisition of multiple Nom case marking	65
2.6.1. Matsuoka's (1998) acquisition study on Japanese MNC	65
2.6.2. Input patterns of MNC in Korean	67
2.6.3. Learnability of MNC	69

2.7. Experiment	71
2.7.1. Experimental design	71
2.7.2. Predictions	74
2.7.3. Results and discussion	76
2.8. Conclusion.....	77

Appendix to Chapter 2

Chapter 3: Scrambling Theories from Acquisitional Perspective

3.0. Introduction	83
3.1. Scrambling.....	85
3.2. Syntactic properties of scrambling	88
3.2.1. Optionality and dual properties.....	89
3.3. Pre-subject scrambling: both A and A-bar properties.....	96
3.4. Syntactic theories of scrambling.....	98
3.4.1. EPP/Focus approach: Miyagawa (2003).....	100
3.4.2. Scrambling feature approach: Grewendorf and Sabel (1999)	103
3.4.3. Case-related approach (Bošković 2004, in press)	106
3.5. Acquisitional implications.....	110

Chapter 4: Learnability Problem in Scrambling

4.0. Introduction	116
4.1. A Learnability puzzle.....	119
4.2. Review of early acquisition studies of scrambling.....	124
4.3. Acc case marker as a cue of scrambling	127
4. 4. Experiment	133

4.4.1. Experimental design.....	133
4.4.2. Method.....	135
4.4.3. Subject	137
4.4.4. Predictions	138
4.4.5. Results.....	139
4.4.6. Discussion.....	141
4.5. Some speculations.....	143
4.6. Implications for the theory of parameter setting.....	144
4.7. Implications for the theory of learnability	146
4.8. Conclusion	148

Appendix to chapter 4

Chapter 5: Concluding Remarks

5.0. Summary.....	153
5.1. Significance of complementing longitudinal production data with experimental studies.....	156
5.2. Optionality vs obligatoriness in language acquisition	157
5.3. Statistical learning used in frequency matching.....	160

References

List of Tables

Chapter 1

Table 1 Verb placement in negatives as a function of tense (from French)

Table 2 Finiteness versus verb placement (from German)

Chapter 2

Table 1 Age of the first occurrence of case markers (Japanese)

Table 2 Percentage of the three particles in adult speech (Japanese)

Table 3 Usage of tense morpheme by children (Japanese)

Table 4 Percentage of the three particles in adult speech (Korean)

Table 5 Frequency of Nom and Past tense in the input (Korean)

Table 6 Rate of Nom/Acc out of potential contexts in adult-to-child speech

Table 7 The first appearance of stative predicates (psych-predicates) and Nom objects

Table 8 Multiple Nom case constructions in child speech

Table 9 Input frequency of psych- predicates with MNC/Nom objects

Table 10 Input frequency of all types of stative predicates and MNC

Chapter 3

Table 1 Typology of scrambling

Table 2 Cross-linguistic survey of scrambling and MNC

Chapter 4

Table 1 Frequency of word order patterns in the child speech

Table 2 Frequency of word order patterns in MJ's mother

Table 3 Frequency of occurrence of the Acc marker in mother's speech

Table 4 Patterns of Object marking in parental input

Table 5	Typology of scrambling and Acc marker
Table 6	Acquisitional implication of scrambling theories

Chapter 5

Table 1	Cross-linguistic variation in the rate of overt subject
Table 2.	The age of adult-rate use of overt subjects

List of Figures

Chapter 2

- Figure 1** Percentage use of the case marker by Hyuck and his mother.
- Figure 2** Percentage use of the case marker by MN and his mother.
- Figure 3** Percent correct production of ASL morphemes, for Stewart as compared with his hearing parents
- Figure 4** Percent correct production of ASL morphemes, for Sarah as compared with her hearing parents

Chapter 5

- Figure 1** Pragmatic differentiation of referential options

Abbreviation Keys

Acc	Accusative
Comp	Complementizer
Dat	Dative
DS	Dative subject
Decl	Declarative
NM	Nominalizer
MNC	Multiple Nom Case
Nom	Nominative
Pst	Past
Q	Question

Chapter 1. The Theory of Language Acquisition

1.0. Introduction

The simple fact that there is no universal language, but only particular languages such as English or Korean, etc, can easily give an appearance that language learning is exhaustively determined by the information arrayed in the language input. Indeed this intuitive idea was in John Locke's mind, when he extended his empiricist view on human knowledge to the knowledge of language. Under the empiricist view that all knowledge comes from our experience through senses, the diversity of language seems to confirm that language learning is purely input-driven, and therefore, there is no such thing as innate knowledge of language.

An assumption that has been developed along this tradition is that the input that the child receives is regular and systematic, displaying distributional properties, and that the child relies on those properties to figure out systems of her target language using general inductive reasoning mechanisms that are already available for other domains of learning. Frequency of linguistic elements has been considered one of the prominent distributional properties of the input that is available for the child to extract information about the language being learned. It was assumed that the more frequently a certain linguistic expression is available to the child, the easier it is for the child to learn it (MacWhiney 1978, Bates and MacWhiney 1982, MacWhiney et al. 1984, Bates and MacWhiney 1987). Intuitively, this simple assumption seems to make sense, when we consider many cases of learning in other domains.

However, a considerable amount of language acquisition research for the last several decades has revealed numerous findings that challenge such an input-driven language learning approach. The discovery starts with Chomsky's (1959) proposal for a new approach to the nature of knowledge of language. Chomsky points out a remarkable fact about language; namely that the

child comes to know abstract properties of language for which there is no corresponding positive evidence. He concludes that the input that the child receives is far from being enough to explain the richness of the output grammar attained by the child. This learnability problem leads him to propose that children are endowed with some prior knowledge of language which enables them to achieve an adult grammar on the basis of limited evidence. Chomsky (1965) conceives the system of prior knowledge comprising universal properties of all human languages, and later, he calls it 'Universal Grammar' (Chomsky 1975). To stress the importance of learnability problem, Chomsky (1965) introduces a notion of 'Explanatory Adequacy' as a goal of theory of language. He argues that theory of language must be able to explain how grammatical properties are actually learnable, given boundary conditions of experience. This idea has established an important perspective that acquisitional data is important empirical evidence that theory of language has to consider.

The theory of universal grammar (UG) takes an increasingly explicit and systematic form, as Chomsky (1981, 1986, and 1995) conceives notions of 'principles and parameters' as two axes that organize the grammar of human language. According to this model of grammar, UG is defined as a set of invariant principles which are common to all human languages, together with parameters which specify permissible options for variation. This theory brings in a new perspective on language acquisition and the theoretical tools which allow a systematic study of language development. Acquisition of language is viewed as a process of setting parameters, rather than figuring out rules of grammar sentence by sentence. Some acquisition studies have supported the Principles and Parameters' (P&P) model based on empirical findings such as children's adherence to some abstract principles of language in the absence of relevant evidence (Thornton 1990, Wexler and Chien 1985).¹

¹ See Thornton (1990) for the evidence for empty category principle, which is assumed to constrain the relation between a gap and a displaced element, and Wexler and Chien (1985) for Binding principles, which determine the coreferential relations between anaphors and their antecedents.

The UG-based approach with the P&P framework has been a dominant view in acquisition research for the last several decades. Since 1990s, however, there has been a revival of an input-driven learning approach in the name of models such as ‘statistical learning model’ (Saffran et al 1996, Newport and Aslin 2000) or ‘usage-based theory of language acquisition’ (Bybee 1985, Goldberg 1995, and Tomasello 2003). Unlike the UG-based approach, these theories assume that input is a rich source of information for identifying grammatical regularities, and argue that children have a remarkable ability to perform complex computations over statistical information displayed in the input, as described below.

Corpus will provide a much richer source of information for identifying grammatical regularities than we have previously thought. (Newport and Aslin 2002)

Infants have access to a powerful mechanism for the computation of statistical properties of the language input. (Saffran, J.R, Aslin, R., and Newport, E. 1996)

This approach has been gaining so much popularity in acquisition research that now the UG-based approach is often referred to as a ‘traditional view’. This thesis is couched in this current state of acquisition research. We evaluate the two views by investigating the acquisition of aspects of grammar in which the role of experience is inevitable. Particularly, we study the acquisition process involved in the learning of syntactic properties that are subject to cross-linguistic variation.

Concerning the acquisition of language particular properties, there is an assumption widely adopted in acquisition research; namely that language input available to the child is fully representative of samples of all the relevant properties possible in the adult grammar. This assumption seems to be found both in the input-driven learning approach and the UG-based learning approach. In the parameter setting model of language acquisition, the child is to set

parameter values on the basis of positive evidence. The notion of ‘trigger data’ expresses the idea that there are sentences in the child’s experience that point directly to the correct setting of parameters (Gibson and Wexler 1994). Lightfoot (1998, 1999) suggests that a robust amount of such data should be available in the input. Some corpus studies seem to support this assumption. For instance, children learning English, which is a language with Wh-movement, a requirement of wh-words to appear in the beginning of the sentence, receive over 40% of wh-questions exhibiting this movement (Newport et al 1977, Valian 1991). Children learning Korean, which allows null pronominal arguments *pro*, are provided with sentences with null subjects at the rate of 50-80% (Kim 2002). Children learning German and Dutch, which require a verb to be in a second position, receive unambiguous verb-second sentences around 30% of the time (Lightfoot 1999).

This thesis investigates some cases which seem to challenge the above assumption. We report an apparent discrepancy between adult grammar and language input available to the child. The following shows such a case of discrepancy between Korean adult grammar and the language input presented to the child, with respect to some language particular properties.

(1) Discrepancy between the adult grammar and the language input

Korean	
Adult grammar	Language Input
$\left\{ \begin{array}{l} +\text{Pro-drop} \\ +\text{Multiple Nom case} \\ +\text{Scrambling} \\ +\text{Dative subject} \\ \dots \end{array} \right\}$	$\left\{ \begin{array}{l} \text{Pro-drop (50-80\%)} \\ \text{Multiple Nom case (10\%)} \\ \text{Scrambling (0.003\%)} \\ \text{Dative subject (0.00\%)} \\ \dots \end{array} \right\}$

Compared to the rate of supply for Wh-movement, Verb-second, or Pro-drop, which runs above the rate of 30% of use, the positive evidence for multiple case marking, scrambling and dative

subjects is drastically rare. This raises a learnability question: how do Korean children learn those language particular properties under the impoverished input?

By investigating the learnability problem, we aim to answer the following three related questions: i) How crucial is the role of input-frequency? ii) Is Principles and Parameters learning model inevitable? If so, what language specific knowledge is required to guarantee the learning? iii) What kind of experience triggers the learning? We ask to what extent distributional properties in the input such as frequency can account for acquisition data. We also consider a UG-based learning approach, which assumes an innate component of language specific knowledge; we ask whether the Principles and Parameters learning model of UG-based learning approach is indispensable, and if so, what language specific knowledge is required. The investigation of the learnability question will lead us to find out what kind of linguistic experience serve as a trigger for the learning.

The domain of language particular properties that will be investigated in this thesis concerns two grammatical phenomena in Korean: Multiple Nominative Case marking (MNC) and scrambling, as illustrated in (2b) and (2d), respectively.

(2) a. Mary-ga John-ul bulesstta

M-Nom J-Acc called

‘Mary called John’.

b. Mary-ga cha-ga pilyohatta

M-Nom car-Nom need

‘Mary needs a car’.

c. Mary-ekey cha-ga pilyohatta.

M-Dat car-Nom need

‘Mary needs a car.’

d. cha-ul_i Mary-ga t_i palasstta

car-Acc M-Nom sold

‘Mary sold a car’.

Example (2a) illustrates a typical sentence in Korean. The subject is marked with overt Nominative case, and the object is marked with overt Accusative case. The two sentences in (2b) and (2d) exhibit some peculiar grammatical properties, when they are compared with (2a). First, consider the example of MNC given in (2b). (2b) has the object marked with Nom case, instead of Acc case, forming multiple instance of Nom case. The subject can appear with Dative case, as shown in (2c). For this reason, this type of sentences with Dative subjects will be investigated along with sentences with MNC. (2d) differs from (2a) in word order. The object in (2d) appears in the sentence initial position, giving rise to the OSV order. This word order permutation is called ‘scrambling’.

We examine the acquisition process of these two grammatical properties using both longitudinal speech data and data from experimental studies. We investigate whether an input property such as frequency accounts for acquisition orders, and if not, what kind of specific syntactic knowledge should be assumed to guarantee the learning. We consider UG-based syntactic analyses proposed to account for the two phenomena, and draw acquisitional implications of each syntactic analysis. We evaluate them by considering whether they make correct predictions on the time course of acquisition that we observe, and further whether each analysis guarantees reliable and plausible positive evidence.

In this introductory chapter, we discuss some previous acquisitional findings with respect to the input-driven learning approach, and review motivations for the UG-based learning approach. Section 1.1 discusses some acquisitional facts regarding the issue of the role of frequency, and shows aspects of a gap between frequency patterns in the input and hypotheses children actually entertain. Section 1.2 questions whether a general inductive reasoning mechanism such as generalization can be a possible mediator for the gap. It will be argued that

even though children seem to use some kind of generalization over input data, they use what is called ‘discriminative generalization’ more prevalently (Fodor and Crowther 2002). We propose that children have to know when to ‘overgeneralize’ (generalize beyond the input) and when to ‘undergeneralize’ (generalize only on a subset of the data). Section 1.3 reviews Chomsky’s (1975) original argument for the poverty of stimulus problem with English Yes-No question formation to reconsider the status of ‘generalization’ in language learning. Section 1.4 takes a look at some well-known errors reported in language acquisition research, and discusses them with respect to the input-driven learning approach. Section 1.5 addresses how the phenomena which are problematic for the input-driven learning approach can be accounted for in the UG-based learning approach. Section 1.6 provides an outline of the dissertation.

1.1. Is Frequency a Determining Factor in Language Development?

Various studies have investigated whether the frequency with which parents use certain constructions or lexical items has indeed an impact on the development of child grammar. Gathercole (1986) study whether the frequency of use of certain linguistic expressions by adult speakers is mirrored in child grammar. Observing that the past participle is more often used in Scottish English than in American English, they test whether the difference between the two dialects in the frequency of the use of the past participle is found in child language. They examine the spontaneous speech data of 3 to 5 year old children, and discover that children in both dialects show sensitivity to the frequency of the participle constructions in the adult speech of their environment. Scottish English learning children used the past participle more often than American learning children, as the adults. This effect of the frequency of use in the child language is interesting in itself, but we do not know whether the difference in the frequency of use has any qualitative significance in linguistic representation. This may be a matter of preference for using a certain construction. What would be more crucial to know is whether the frequency affects the order of acquisition.

One surprising fact that has been discovered in early language acquisition research is that 62% of total utterances in the input are questions and commands, and only 30% are simple declaratives, even though simple declarative sentences are the first kind of sentences that children produce (Newport et al. 1977). This fact has been interpreted as an indication that the frequency itself does not play a decisive role in determining the order of grammar development. However, one could consider the possibility that somehow the psychological state of the child in the early stage is such that they are more inclined to make statements or descriptions rather than to ask questions or to make commands.

Snyder and Stromswold (1997) provide a case in which such a language external consideration does not interfere. Their study includes the acquisition of two constructions which express more or less the same meaning: double object datives (*John will give Mary the book*) and *to*-datives (*John will give the book to Mary*). Given that the meaning conveyed by the two constructions is almost the same, the input-driven approach would predict that the frequency of the constructions in the input will determine the acquisition order of the two constructions. They analyze 12 children's spontaneous speech in CHILDES (MacWhiney 1996), and find that all children acquire double object datives significantly earlier than *to*-datives. They consider the possibility that the order is an effect of frequency in the input rather than acquisition of syntactic knowledge. In order to evaluate this possibility, they examine whether children who hear more *to*-datives acquire *to*-datives earlier than children who hear few *to*-datives. It turns out that there was no significant correlation ($p > .10$) between the relative frequency of *to*-datives in the adult speech and the age of acquisition of *to*-datives by the children. This fact suggests that frequency does not shape the order of grammar development.

Dutch children's use of non-finite verbs and word order patterns in their early stage provides another argument for the above conclusion. Wijnen and Verrips (1998) examine Dutch children's longitudinal speech data, and report that some children produce *infinitives* (non-finite verbs) in finite contexts for nearly 100% of the verbs during the first month after the onset of

combinatorial speech. Children's use of non-finite forms in finite contexts has been reported in a variety of languages under the name of 'root infinitive' or 'optional infinitive' phenomenon, and it has generated a considerable amount of research concerning the properties of the phenomenon. The following shows samples from Dutch children.²

(3) a. thee drinken

tea drink-INF 'drink(ing) tea' (Niek, 2;7)

b. die helemaal kapot maken

that + one altogether broken make-INF (Niek 2;7)

'smash that one.'

Dutch data is significant in that the infinitive verb is inflected with the suffix '-en', not a bare stem as in English. This means that the infinitives cannot be analyzed as simply leaving out inflections, as one could assume for English. Here, our concern is the question of whether this is attributable to frequency in the input patterns in any sense.

Wijnen (2001) explores the possibility that the early predominance of infinitival forms is related to the use of such forms in the language input. They find that infinitive utterances in the input amount to only 10% or less, and further that even auxiliary-plus infinitive main verb constructions, which in terms of word orders are maximally similar to infinitives, constitute at most 30% of all verb-containing utterances.³ This is a clear indication that children's use of non-finite verbal forms in the early stage cannot be due to the frequency of those forms in the adult speech. There seems to be a gap between what the children hypothesize and what is expected if the frequencies displayed in the input play a major role in shaping the process of language development.

² Wijnen and Verrips (1998) note that forms similar to imperative are often observed, but in many cases, imperatives were unproductive routines and those imperatives were not included.

³ When a sentence contains an auxiliary, the main verb appears in the final position.

Word order patterns in child Dutch point in the same direction. Dutch is a verb second language (V2), a class that includes German, Danish, Swedish, Yiddish, Icelandic, among others. The characteristic syntactic phenomenon of V2 is that the finite verb is always the second constituent in matrix clauses, and some other constituent such as subject, object, or an adverbial maximal projection appears in the first position (SVO, OVS, XP VSO, etc.). The position of the finite verb in matrix clauses is assumed to be the position of a complementizer. The evidence for this assumption comes from the complementary distribution of finite verbal elements and overt complementizers; when an overt complementizer is introduced in the subordinate clause, the finite verb appears in the final position, indicating that the underlying basic word order is head-final OV.

Corpus studies (Yang 2000, Lightfoot 1989, 1999) on German and Dutch report that the VO pattern constitutes more than 90% of the child-directed speech (SVO: 70%, XP VS (O): 26%, OVS: 7%). Given this overwhelming VO pattern, if the child predominantly relies on frequency in the input, it is expected that she will start by producing the VO pattern rather than then OV pattern even if the OV pattern is the basic word order. In contrast to this expectation, Wijnen and Verrips (1988: 646) report that 87%-95% of the early two-word utterances are OV patterns with objects preceding verbs.⁴ This data is striking in that even the most easily noticeable regularity such as that of word order does not directly influence the learning process. The issue of why Dutch children predominantly used the OV pattern in their early stages, in spite of its lower frequency in the input, is discussed in section 1.4.

Bates and MacWhiney (1987) develop a refined notion of frequency called 'validity', assuming that statistical properties such as frequency and regularities account for the nature of grammar, and further play a major role in the order of acquisition. The notion of 'cue validity' is defined over 'availability' (how often a cue is available) and 'reliability' (how often a cue leads to a correct conclusion when it is used) of some linguistic element. Two grammatical properties are

⁴ Wijnen and Verrips (1998) note that most of the utterances with the OV order contain non-finite verbs.

recognized as important regularities that are normally found in language: word order and case markers. It was argued that languages use those two kinds of ground plans to express grammatical relations. English was considered as a word order language, while Turkish belongs to the group of case marker languages. This leads to a hypothesis that the task of the child as a learner is to recognize what ground plan is used in her language. For languages which use a mixture of the two ground plans, it was hypothesized that the child uses the cue with a higher validity. A flurry of studies was done in various languages to evaluate this approach (Slobin and Bever 1982 for Turkish, Slobin 1971 for Russian, Bowerman 1973 for Finnish, Miyahara 1973 and Morikawa 1989 for Japanese)

Chung (1994) takes this line of approach to account for Korean children's acquisition of grammatical relations. He observes in his experimental study that some Korean children fail to understand scrambling sentences with OSV word order, incorrectly choosing basic word order strategy (SOV) to interpret the sentences. Chung found that the cue validity of word order was .72, whereas that of case marker was .32 (Chung 1994:137).⁵ Considering the difference in the cue validity, he interprets the results as indicating that children use word order because word order has a higher validity than case markers in Korean, and that the cue with lower validity is learned

⁵ The cue validity of word order and that of case marker (Nom) are calculated in the following way (Chung 1994:135).

(i) word order

Availability: $\frac{\text{Total number of occurrences of first NP}}{\text{Total number of utterances}}$

Reliability: $\frac{\text{Total number of subjects that appear in first position}}{\text{Total number of occurrence of the first NP}}$

Cue validity: availability (1.0) x reliability (.72) = .72

(ii) Nom case

Availability: $\frac{\text{Total number of occurrences of } ka}{\text{Total number of utterances}} =$

Reliability: $\frac{\text{Total number of } ka \text{ used for subject}}{\text{Total number of } ka \text{ used for subject and object}}$

Cue validity: availability (.34) x reliability (.94) = .32

Given the fact that Nom case is obligatory, this low cue validity of Nom may seem surprising. According to Chung's calculation, the low validity is due to the low availability of Nom case. I assume that that has to do with the frequent use of null subjects, which is around 50-80% out of total utterances (Kim 2002).

later. Based on this, Chung concludes that children note what is constant and what is variable in the input, and uses the difference in the pattern to learn the grammatical system of their language.

The above frequency-based analysis of children's failure with OSV scrambled sentences leaves several questions unexplained. As we will see in chapter 2, various longitudinal acquisition studies on case markers report that Korean children normally acquire Nom case markers around the onset of two word utterances, indicating that Nom case marker is acquired no later than basic word order. This means that the validity defined by frequency cannot explain the early acquisition of Nom case markers, suggesting that children's failure with OSV utterances cannot be ascribed to the low cue validity of case markers in general. This calls for an alternative account. We take up this issue in chapter 4. There we examine the grammatical properties of scrambling closely, and discuss problems with this analysis.

To summarize, the series of studies discussed above indicates that the frequency of certain forms of lexical items, certain word orders, or certain constructions does not determine the order of acquisition, suggesting that the learning process that the child goes through is not shaped by what is obvious on the surface. There is a mismatch between frequency patterns found in the input and what the child actually opts to entertain in language development. This leads us in the direction that there must be some kind of mediator between the input and the hypotheses that the child entertains in the language learning task.

1.2. Could General Inductive Reasoning Reconcile the Mismatch?

Could that mediator be some kind of general cognitive reasoning such as generalization, analogy, or association? Since those tools are used for other domains of knowledge, it would be parsimonious to assume that they are also employed for the task of language learning. When we look at children's spontaneous speech data, we often find cases that may be analyzed as applications of generalization. For instance, the child fills in different lexical items in some

combinatorial slots. They can replace different nouns in adjective-noun frames. Being exposed to some verbal phrases, they can insert different nouns or verbs in verb-object slots.

However, on a closer look, we find more overwhelming cases in which children go beyond this lexical level of simple substitution operations. They pick up knowledge that allows insertion of the clausal unit in place of a simple nominal phrase even if examples of such types of sentences are not likely to appear in the input. They also come to know the unboundedness of wh-movement by hearing wh-questions in a single clause context (e.g. *what did Daddy eat?*). The possibility of hearing even 3 levels of embeddings would be close to zero. However, the child comes to know that she can apply wh-movement unboundedly, not just in root clauses, but also in *n*-number of clauses (e.g. *What did you say that Mommy thought that Mary ate?*). These cases demonstrate that the scope of generalization that the child makes goes far beyond what is available in the input. This means that the child as a learner has to make generalizations for which there is no corresponding evidence. Many more aspects of language exhibit unboundedness, requiring the child to make generalizations in the absence of positive evidence. Fodor and Crowther (2002) call this situation a case of poverty of positive stimulus. This is certainly an important aspect of language acquisition which calls for an explanation in any language learning theory.

Could we resolve the problem of poverty of positive stimulus if we assume that the child somehow makes generalizations blindly? In other words, could the mechanism proceed so that whatever happens in root contexts always happens in the non-root contexts, salvaging the problem of poverty of positive stimulus? It does not seem to be the case. Natural languages exhibit several phenomena which are sensitive to the root vs non-root distinction. For instance, in many languages, the requirement that verbs appear in the second position is only imposed in main clauses, not in embedded clauses. Also, French allows wh-phrases to stay in-situ only in root clauses, but not in embedded clauses (Bošković 2000). The internal grammar seems to be

organized in such a way that certain generalizations are blocked. This means that the child cannot just apply them blindly. She should know the upper bound of generalization in some cases.

Not only does language learning involve ‘overgeneralizing’ beyond what is given in the input, but it also requires the exact opposite situation. It sometimes demands ‘undergeneralizing’, in the sense that it requires the learner *not* to make a generalization that would be perfectly reasonable given the data provided. The following two examples illustrate celebrated cases of such a situation: *wanna*-contraction (4) and *that*-trace effect (5), respectively (Thornton 1990 and Crain 1991).

(4) a. Who do you want to help?

b. Who do you wanna help?

c. Who do you want to help you?

d. *Who do you wanna help you?

(5) a. Who do you think Mary will call?

b. Who do you think that Mary will call?

c. Who you think will call John?

d. *Who do you think that will call John?

As indicated, each paradigm contains one instance marked as ungrammatical. If one makes the generalization of *wanna*-contraction based on (4a) through (4c), (4d) is expected to be grammatical along with (4c). The same line of reasoning holds for the paradigm for *that*-trace effect in (5). Given the well-formedness of (5a) through (5c), one would conclude that the complementizer *that* is optional, and expect (5d) to be grammatical. These paradigms illustrate so-called ‘constraints’. Natural languages are known to exhibit some limited number of grammatical constraints. They act as filters, closing off certain routes of grammar formation,

creating gaps in the paradigm on the surface (constraints on form) or blocking certain interpretations (constraints on meaning). The above two cases illustrate the first type, a constraint on form.

Does the child ever incorrectly try out these forbidden options, violating the constraints? Thornton (1990) conducted experiments to investigate this question, and demonstrated that the child obeys the constraints in both cases, even if there were other kinds of non-adult forms. How does the learner know about the gaps in the paradigm? Is it possible that they could obtain this information from the input? It has been an established fact that the child does not receive negative information (the information on what is not possible in the target grammar) in any systematic way (Brown 1973, Marcus 1993). This means that there is no resource available in the input that can tell the child not to make those particular generalizations, but somehow they obtain this knowledge.

The above illustration shows that language learning requires knowing certain constraints on generalizing in the absence of evidence, a situation called the *poverty of negative stimulus*. Could we rely on some sort of *conservatism* as an explanation for this poverty of negative stimulus? One may consider the possibility that the child is conservative in that she restricts her hypotheses to the data that she observes. This may work for this case, but it will get us into trouble for the cases in which the child does have to make generalizations beyond what is in the input.

To summarize, the task of language learning presents two situations with respect to the kinds of generalization that it employs; it demands generalization to go beyond the information in the input, and at the same time it requires generalizations to be made only based on some subset of input data, banning full use of the information in the input data. In other words, it requires the child to make discriminative generalizations. This means that we need to specify conditions for when different kinds of generalization should apply. Obviously, this complicates the learning

theory, taking us further away from our original motivation of considering the more parsimonious approach.

1.3. Is it Generalization at all?

The breakdown of simple generalization with the two sides of the poverty of the stimulus makes generalization as a main tool for language learning questionable. Then, one could wonder whether language learning still involves generalization, but with upper bounds and lower bounds and no other mechanisms. However, the way language works seems more complex than that. The basic rules that operate in language organization do not seem to be the kind of generalization that we normally assume in data processing. Chomsky (1975), taking the example of Yes-No question formation in English, succinctly shows that the rules involved are not the kind of generalization expected in general inductive reasoning.

Let us review the gist of the argument. Yes/No questions in English are formed by inverting the auxiliary with the subject, as in (6). There are many possible hypotheses capable of generating the surface strings. Among these hypotheses, (7a) seems to involve a particularly simple and elementary operation. It only makes reference to linear order, hence, it is a 'structure independent rule'. On the child's part, she only needs to scan down the linear string, and when she hits the auxiliary, she simply needs to invert it. However, this hypothesis generates incorrect results when more complicated data are involved, as the one given in (8a). In order to generate the correct Yes-No question form (8b), the knowledge of a 'structure dependent rule' is required, as in (7c). The child needs to look at the structural representation of the sentence to pick out the structurally higher INFL element.

(6) a. Will Mary come?

b. Is Alex singing a song?

c. Has Robin finished reading?

- (7) a. front the first aux following the first NP
b. front the aux whose position in the sentence is a prime number
c. front the matrix I to C
- (8) a. The man who is beating a donkey is mean.
b. Is the man who is beating a donkey mean?
c. * Is the man who beating a donkey is mean?

Now, compare the two hypotheses from the learning point of view. The structure independent rule is simple and elementary. If the types of the Yes-No questions that the child often hears are those in (6), the child has every reason to adopt the structure independent rule. The structure independent rule can analyze almost all instances of Yes-No questions that are available to the child, and the rule itself requires only a simple operation. Then, why would the child bother to discard this option, and go for a complicated structure dependent rule? Or alternatively, does the child first attempt the simple and elementary structure independent rule, and later change to the complicated structure dependent rule? Chomsky conjectures that the child would never entertain the structure independent rule at all, even if it is highly likely from the point of general reasoning. He suggests that structural dependence is an innate property of language.

Crain and Nakayama (1987) carry out an experiment to test Chomsky's hypothesis on the learning of Yes-No questions. They find that some children succeed in making correct Yes/No questions, and some children fail. However, none of the children who fail producing grammatical Yes/No questions violate the structure dependent rule, showing that the child chooses a rule that is very unlikely, without entertaining the simpler option. This fact suggests that the hypotheses that the child formulates in grammar development make reference to hierarchical structure of the sentence, instead of generalizations that can be drawn from general inductive reasoning. This

constitutes another strong argument for the conclusion from the poverty of stimulus argument, that aspects of knowledge of language are innate.

Recently, this argument has been put under a close scrutiny by Pullum and Scholz (2002). They claim that there is enough positive evidence of the type (8b) available to the child, challenging the basis of the poverty of stimulus argument.⁶ However, the empirical data that they rely on is from an adult magazine the *Wall Street Journal* corpus, which contains structures very unlikely to be available to the child.⁷ Legate and Yang (2002), pointing out the absurdity of using the adult magazine as a possible source of positive evidence for the child, examine child-directed speech taken from the Nina corpus from CHILDES. The corpus contains 46,499 sentences of child-directed adult speech, out of which 20,651 are questions. They report that among those there were only 6 instances of Yes/No questions of the type in (8b).⁸ This frequency corresponds to 0.068% (taking all yes/no questions as the denominator), which seems too low to be useful to the child. They also find a similar proportion in Adam's corpus. This fact shows that Chomsky's original argument from the poverty of the stimulus still holds.

So far we have discussed the possibility of bridging the gap between the input and the hypotheses the child entertains with the mechanism of generalization used in other forms of learning. We showed that generalization faces not only the problem of poverty of positive stimulus but also the problem of the poverty of negative stimulus. Further, rules that operate on language organization do not seem to be the kind of generalization that we normally adopt in data-processing. This situation invites us to assume that the gap cannot be mediated by inductive

⁶ Pullum and Scholz (2002) estimate that 0.1-1.0% of all adult input sentences show the evidence for (8b) and assume that that rate is enough.

⁷ Some instances of the type (8b) they found from the *Wall Street Journal* are shown below ((Pullum and Scholz 2002:42)

- (i) a. How fundamental are the changes these events portend?
- b. Is a young professional who lives in a bachelor condo as much a part of the middle class as a family in the suburbs?
- c. Is what I'm doing in the shareholders' best interest?

⁸ Some examples are shown below (Legate and Yang 2002).

- (i) a. Where's the little red duck that Nonna sent you? (NINA02.CHA)
- b. Where are the kitty cats that Frank sent you? (NINA03. CHA)
- c. What is the animal that says cockadoodledoo? (NINA04.CHA).

reasoning such as generalization over input data, and that language learning requires a different kind of mechanism. In the following sections, we look at some errors that are produced by the child to see what these facts suggest.

1.4. Children's Errors

One striking fact that we see in spontaneous speech data is that children make relatively few errors. Significantly, those errors found are not restricted to an individual child, but rather they are observed across children. These two observations are already an alarm call that an input-driven learning approach may not be on the right track; since the input that the individual children receive varies considerably we would expect heterogeneous errors. Various studies have shown that even the errors reveal some intricate patterns that cannot be ascribed to input data. Here we take a look at a couple of error types: Overgeneralizations, Root infinitives, Medial wh-questions in long distance questions and Resumptive pronouns in relative clauses.

First, we consider Overgeneralization errors. It is a well-known observation that English learning children overgeneralize past tense suffix *-ed* and the plural suffix *-s*, producing *goed* and *foots*, for example. The child is applying the generalization for pluralization or past tense formation across the board, without having learned that there are some irregular verbs or nouns. This overgeneralization fact clearly indicates that children are applying rules, instead of just imitating adult speech. Pinker (1994) points out that what the child does is not simply generalization. He refers to Stromswold's (1990) finding that children overregularize main verbs like *have*, *do*, *be*, just as they overgeneralize other main verbs, but they do not overregularize the same verbs when they are used as auxiliaries. For instance, children produce utterances as in (9), but not the ones in (10).

(9) a. I *doed* it (I *did* it)

b. I haved it (I had it)

(10) *a. Doed you come? (Did you come?)

*b. I haved eaten. (I had eaten.)

The above pattern shows that children's rule operation is distinct in that it is sensitive to syntactic environments where the operation is applied, rather than being a simple blind application of generalization.

Another type of error known as the Root infinitive phenomenon displays a similar property. As briefly mentioned in section 1.1, this error has been found in various languages. The following are some samples of the error (Guasti 2002: 128).

- | | |
|--|---------|
| (11) a. Cromer wear glasses. (Eve, 2;0) | English |
| b. Hun sove. (Jens, 2;0) | Danish |
| she sleep-INF 'She sleep.' | |
| c. Dormir petit bébé (Daniel, 1;11) | French |
| sleep-INF little baby 'Little baby sleep.' | |
| d. Earst kleine boekje lezen (Hein, 2;6) | Dutch |
| first little book read-INF ' | |
| 'First (I/we) read little book.' | |
| e. S[ch]okolade holen (Andreas, 2;1) | German |
| chocolate get | |

The Root infinitive phenomenon refers to cases in which the child uses infinitive verbs instead of finite verbs in main clauses. The nature of the error has become better understood through studies of a wide variety of languages including English, Danish, Dutch, German, Russian, and Swedish,

which have overt suffixes for infinitive forms (See Platzack 1992, Wexler 1994, Haegeman 1995, and Hamann and Plunkette 1998). The cross-linguistic facts show that the uninflected verbs that the child produces are not just cases of omission of inflection, nor repetition of frequent forms of the verbs that they hear, as pointed out previously.

There are two intriguing properties of this error that are relevant here. One is its cross-linguistic distribution. It has been reported that the error is found only in non-pro-drop languages, in which null subjects are not a grammatical option, but do not appear in pro-drop languages such as Catalan, Spanish or Italian (see Guasti 2002). If the error were the result of some common generalization, we would expect the error to show up in all languages. The fact that the phenomenon is restricted to non-pro-drop languages indicates that there is some grammatical relation between non-pro-drop and root infinitives that cannot be accounted for by a non-linguistic factor.

Another property, probably a more intriguing one, is the systematic distribution of the infinitives with respect to other constituents within a sentence. Languages with so-called verb movement illustrate this point clearly. It has been known that languages vary as to where verbs can appear, and further that within a given language, the distinction of finite vs non-finite makes a difference as to whether the verb can appear in a certain position or not. The finite verb is associated with a higher position in the structure, and non-finite verbs with a lower position. This correlation has been analyzed in terms of the presence or absence of verb movement. French is known to have finite verb movement to T, and Dutch or German are known to have finite verb movement to C in matrix clauses, while infinitive verbs in both types of languages do not need to move, and stay in their base-generated positions. For instance, in French this contrast between finite verbs and non-finite verbs in terms of the position that is particularly observable with respect to negation *pas*. When the verb is infinitival, it appears after negation *pas*. However, when it is finite, it appears before negation *pas*, as in (12).

(12) Le chat (ne) chasse pas le chien.

The cat chases not the dog

‘The cat does not chase the dog.’

Now, going back to children’s root infinitives, if it were an error from generalization, we might expect them to use the infinitival verbs indiscriminately. In other words, we expect that they use them both before and after negation in this case. Pierce (1992) reports that almost all root infinitives in four children’s data follow negation. The following data are from one child in her study.

(13) Table 1. Verb placement in negatives as a function of tense (from French)

(Pierce 1992)

	+Finite	-Finite
Verb-Neg	68	-
Neg-Verb	3	82

(14) Data showing that +/- finite determines location of Negation

a. pas manger (Natalie 1-9-3)

not eat [-finite]

b. veux pas lolo

want [+finite] not water

The data show that the child obeys the correspondence between a certain position of the verb and finiteness.

The same correspondence is observed in languages where the finite verb has to move to C, putting the verb in the second position in the sentence. In these languages, finite verbs appear in C, while non-finite verbs stay in their base-generated position. The distribution in table 2 shows that the German child treats finite verbs and infinitives differently, putting only finite verbs in the second position, while keeping the infinitives in the base-generated position.

(15) Table 2. Finiteness versus verb placement (from German)

(Poeppel and Wexler 1993)

	+Finite	-Finite
V2	197	6
Verb-final	11	37

This correspondence again reveals that the child is sensitive to fairly abstract relations between finiteness and the need for movement. The nature of the error is still under debate, which we will not go into here. An important point that we want to draw from the fact is that children's errors are not simply the errors of generalization made by not considering some irregular items, but they show some grammatical aspect to an extent that they display interactions with other constituents in the sentence. This aspect cannot be explained by general inductive reasoning because it has to invoke language specific information concerning the abstract relations between linguistic elements.

Another type of error we are about to discuss poses probably the most puzzling problem for inductive reasoning approaches to language learning. It has been noted that some errors that children make turn out to fall under possible options in non-target adult grammars. Thornton (1994), eliciting long-distance questions from English learning children (2;10-5;5), finds that some English speaking children make the error of producing an additional wh-word in the intermediate positions. Some examples of those 'medial wh-questions' are given below.

- (16) a. What do you think *who*'s in there, really, really really? (AM 4;6)
 b. What do you think *what* Cookie Monster eats? (KM 5;5)
 c. What do you think *what* the baby drinks? (MA 3;3)

A striking fact is that there are languages in which this error pattern is actually a grammatical option. For instance, some southern dialect of German allows the following wh-questions.

- (17) Wer glaubst du wer nach Hause geht?

Who think you who to home goes

'who do you think goes home?'

There is another area in which children seem to make the error of inserting an extra word in the sentence. That is relative clauses. Since Labelle's (1990) finding in French, various studies have shown that some children produce resumptive pronouns in relative clauses where adult grammar requires a gap. This kind of error has been attested in various languages: Russian (Bar-Shalom 2001), Spanish and English (Pérez-Leroux 1995), Serbo-Croatian (Goodluck and Stojanovic 1996), Korean (Cho 1999). Some examples are shown below.

- | | |
|--|---------|
| (18) a. celle-là que le lui montre un dessin | French |
| b. the one that he lifted <u>it</u> | English |
| c. vot etot kotoryj kniga upala nanjejo. | Russian |
| the one which book fell on him | |
| d. thokki-ka twayci-lul minun twayci. | Korean |
| rabbit-Nom pig-Acc push pig | |
| 'the pig that a rabbit pushes' | |

English does not allow resumptives in a context such as in (18b), but has certain context in which resumptives become acceptable. For instance, consider (19b). (19b) contains a so-called ‘island’ from which a gap cannot relate to the displaced element any more. The bracket indicates the boundary. The sentence with a gap is ungrammatical, but with the use of resumptive pronoun, it becomes acceptable.⁹ The contrast with (19a) shows that resumptives in English are allowed only in a context where a gap is not possible.

(19) a. *The dress that I bought it was expensive.

b. I’d like to meet the linguist that Mary couldn’t remember [whether she had seen him before.]

Interestingly, it was found that the use of resumptive pronouns as in (18) is not unique to child language. Hebrew, Irish, Palestinian Arabic, Welsh, and Brazilian Portuguese, all allow resumptive relatives in contexts where a gap is also possible. Some examples are shown in (20). Sells (1984) refers to English type resumptives as ‘intrusive resumptives’ and the second type as ‘true resumptives’.

(20) a. ha-simla Se kaniti - /ota yayta yekara Hebrew

the-dress Op I-bought _ /it was expensive

‘The dress that I bought was expensive.’

b. o homem que eu vi _ -/ele Brazilian Portuguese

the man that I saw _ /him

‘the man that I saw..’

⁹ Many native speakers find that the use of a resumptive pronoun in the island does not make the sentence fully grammatical, but makes it less unacceptable (p.c. with William Snyder).

In addition to the syntactic distributions of the resumptives, Sells points out another difference between the two types of resumptives. He shows that there is a difference in the interpretations that resumptives allow. He argues that only true resumptives allow bound variable reading, but intrusive resumptives have only referential readings. The following data illustrate the point.

- (21) a. *I'd like to meet every linguist that Mary couldn't remember whether she had seen him before.
 b. I'd like to meet the linguist that Mary couldn't remember whether she had seen *(him) before (=19b)

In both (21a) and (21b), the resumptive pronoun is in an island. (21a) minimally differs from (21b), in that the head noun is a quantified NP *every linguist*, instead of *the linguist*. In (21a), the resumptive cannot save the sentence. (21b), however, the resumptive pronoun *him* succeeds in saving the derivation. Sells' reasoning on the contrast between the ungrammatical (21a) and the grammatical (21b) with the resumptive is the following. The resumptive pronoun *him* in (21b) happens to pick out some individual, and *the linguist* also happens to pick out the same individual. In other words, there is no actual binding here, but rather the interpretation arises by some kind of accidental coreference. On the other hand, the same mechanism used in (21b) is not available for the resumptive pronoun in (21a) because *every linguist*, being a quantified expression, can only participate in bound variable reading. This leads Sells to conclude that the intrusive resumptive pronoun does not allow bound-variable reading.¹⁰ In contrast to intrusive resumptive pronouns, true resumptives can be interpreted as bound variables, as shown in (22).

¹⁰ Sells' (1984) characterization of the semantics of intrusive resumptives cannot accommodate the following example.

(i) I'd like to meet the only linguist that Mary couldn't remember whether she had seen *(him) before. In (i), the pronoun 'him' can be construed as 'the only linguist', even though the antecedent is a quantifying expression, indicating that the intrusive resumptives can be bound variables. This is problematic for Sells' argument that intrusive resumptives can only be referential. One way to account for

(22) kol gever se dina xosevet se ___ / hu oheve et rina...

Hebrew

every man_i that Dina thinks that ___/he_i loves Rina

'Every man_i that Dina thinks t_i likes Rina...'

The resumptive pronoun *hu* in (22) is allowed even when the relative head is a quantified expression *every man*. Kang (2003), drawing on main characteristics of resumptives, investigates whether English speaking children's resumptives have the same properties as those of true resumptives. She finds that some children do accept resumptives in non-island contexts, and also assign bound variable reading, as in true resumptive languages.

The medial wh-phrase error and the resumptive error show that some children can entertain hypotheses that are not grammatical in their target grammars. Interestingly, however, those errors are found to be grammatical options in other languages. If they are simply errors from inductive reasoning, there is no reason why the child chooses the options found in other adult languages, out of infinitely many possible errors. This fact suggests that the incorrect hypotheses the child entertains seem to be restricted to the boundary of possible human languages.

To summarize, we discussed two important aspects about children's grammar development by observing their error patterns. One aspect was that the errors show grammatical interaction with other constituents in the sentences, obeying a grammatical constraint which regulates the forms of the verbs (finite vs non-finite) and the position in which they can occur. The second important aspect was that the errors exemplify one of the options possible in other adult grammars. These two properties cannot be explained by the mechanism of generalization because they require grammatical information that is not available in the input. There is no basis on which the child can make a generalization. This leads us to assume that the child must have access to extra sources of knowledge that are specific to language in their hypothesis formation.

this varied range of semantics could be that intrusive resumptives are also bound variables, but they range over only a small set of individuals for some pragmatic reason. I thank Yael Sharvit for this point.

1.5. UG-based Approach

The major points of our discussion so far can be summarized as below.

- (23) a. Frequency in input does not determine order of acquisition.
- b. Language learning requires ‘overgeneralizing’ with respect to what is provided in the input.
- c. Language learning requires ‘undergeneralizing’ with respect to what is provided in the input.
- d. Rules that operate on language may not be the generalizations as normally assumed in data-processing.
- e. Errors which children make exhibit grammatical properties that require language specific information, and they are confined to the boundaries of possible human languages.

The above points lead to the conclusion that language learning requires more than what can be drawn from the input by the means of inductive reasoning such as generalization.

To explain the puzzle discussed so far, Chomsky (1965, 1975) proposes a theory of *Universal Grammar*, in which he assumes that the child is born with universal properties about language. In contrast to the empiricist view that our mind at birth is a *tabula rasa* ‘blank slate’, and that all knowledge is the result of experience, Chomsky holds a nativist view to language that our mind has already a richly structured system of knowledge about language even before it is exposed to experience. In this sense, Universal Grammar (UG henceforth) is the theory of the *initial state* S_0 of the mind. Chomsky assumes that language learning can be seen as the function of the initial state S_0 mapping the input to a final state S_f through some transitional states, as in (24).

$$(24) S_f = S_0 (\text{Input})$$

UG consists of common properties that are found in all human languages. In early Generative grammar, X-bar theory for phrase structure rules that are used to build well-formed phrases was conceived as part of UG, along with the general format and constraints of the grammar. Particular grammars are assumed to be systems of language specific rules that are derived by transformation that changes the surface form of the sentence. For instance, rules such as wh-movement or I-to-C movement that is found English, but not in Korean, were analyzed as the result of the application of transformations. In this assumption, the role of the learner is to induce specific rule systems on the basis of experience with the guidelines by UG.

Principles and Parameters Model (Chomsky 1981, 1986, 1995)

An important change in perspective was made in the early 1980s. Extensive comparative studies showed that languages share considerable properties, in spite of apparent differences on the surface. The efforts of incorporating such findings into the linguistic theory shaped the idea that previous construction specific rules can be subsumed under a small number of overarching principles, allowing the possibility that language specific phenomena are derivable from parameterizing the principles. This approach takes on a concrete form in the proposal known as Principles and Parameters (P&P) in Chomsky (1981).

In this model, UG consists of a finite number of principles and parameters. Principles are universal properties of all languages, defining possible human language. Parameters are points of variation, circumscribing the range of possible language variation. Linguistic differences across languages are reduced to a small number of parameters. A particular grammar is an instantiation of UG with parameters specified in one of the permissible ways. There are no construction specific rules. The apparent language specific rules are created by the interaction between invariant principles and the settings of the parameters. Consequently, clusters of properties which look unrelated on the surface can be derived by the same parameter.

This shift in linguistic theory brought a new way to look at language acquisition. Language acquisition started to be viewed as a process of fixing parameters on the basis of primary linguistic data, in addition to learning the lexicon. The process becomes a matter of selecting options out of the already given inventory, rather than a matter of figuring out rules for each construction. This view soon was hailed as a convincing account for some major characteristics of language acquisition. The rarity of errors, relative rapidity, and effortless mastery all seemed to follow more naturally under P&P model, given that the role of the learner is just to pick correct settings of the parameters. This vastly narrows down the search space on the part of the learner. Further, the error types that were discussed in the previous sections receive natural explanations under the P&P model. The puzzling fact that some children incorrectly entertain hypotheses used in other adult grammars can be explained by the assumption that they somehow choose a wrong setting of the relevant parameter that were already available to them as a part of UG.

1.6. Outline of the Dissertation

In the introductory chapter, we discussed some acquisition facts with respect to the input-driven learning approach, particularly focusing on the issue of whether an input property such as frequency determines the shape of language development. The acquisitional findings suggested that input frequency does not determine the acquisition process that children go through, leading us to consider the UG-based learning approach. From chapter 2 through 4, we study the nature of the acquisition process involved in the learning of two language particular properties: multiple Nom case marking and scrambling. We will find a surprising fact that for both properties, the acquisition data show remarkable frequency matching effects, giving us the appearance that language development is driven by input-matching. However, we will see that such data are not the evidence for the input-driven learning approach. New findings in our experimental studies

will show that the acquisition of language particular syntactic properties cannot be explained without the assumptions of language particular innate knowledge such as principles and parameters.

Chapter 2 investigates the acquisition of multiple Nom case marking (MNC). A peculiarity of MNC sentences is that the object appears with Nom case along with the subject, instead of Acc case. We observe from spontaneous speech data that children do not produce MNC sentences even after they have learned Nom case as a marker for the subject and the predicates with the Nom marked objects.

Investigating the nature of the absence of MNC in the child speech, we consider two UG-based syntactic analyses of MNC, which treat multiplicity of Nom case as an additional property that distinguishes MNC from other types of sentences, in addition to the peculiarity of the Nom marked object. The two syntactic analyses of MNC are considered are Default Nom approach by Fukui and Takano (1998) and Multiple feature checking approach by Ura (1996, 1999, 2000). The Default case marking analysis takes multiple Nom as a result of all instance of Nom case in Korean being the default case, while the Multiple feature checking analysis holds that Nom case in Korean is a structural case, and multiple Nom is a result of the positive setting of a parameter, which allows multiple checking of the structural Nom case feature.

Besides the UG-based syntactic theories of MNC, we also consider whether the apparent absence of MNC in the child speech is related to input frequency patterns, rather than the lack of a property required for MNC, as predicted by the two syntactic theories of MNC. We observe that the input provides rather few instances of multiple Nom case utterances, suggesting the possibility that the lack of MNC in children's speech may not necessarily indicates the lack of MNC in their grammar.

In order to determine whether the lack of MNC in the production data is indeed the lack of a certain grammatical knowledge of MNC such as multiplicity of Nom, or whether they actually have the relevant knowledge of MNC, but simply did not produce it, assimilating to their

adult speech patterns, we conduct an experiment. In the experiment, we examine the acquisition of single Nom and multiple Nom. We compare the acquisition of ‘Dative subject’ (DS) sentences with that of MNC sentences as because DS sentences only differ from MNC with respect to the property of multiple Nom. We observe from the adult speech that the input contains virtually no examples of DS sentences while there were some instances of MNC. If frequency determines the order of acquisition, this predicts that DS sentences will be acquired later than MNC sentences because the one with more frequency is expected to be acquired earlier. The two syntactic theories also make predictions distinct from the frequency-based learning approach. The Default case theory argues that the property of MNC is a result of all instances of Nom case in Korean being a default case. The multiplicity of Nom case is deduced from the nature of single Nom case. This means that there is no additional property that the child has to learn, apart from the predicate types and the requirement of marking the object with Nom. This predicts that MNC is likely to be acquired around the same time as single Nom, DS sentences. On the other hand, the Multiple feature checking approach argues that MNC requires the learning of an additional property. MNC becomes available with a positive setting of multiple feature checking parameter. This predicts that MNC will be acquired later than DS sentences because multiple Nom case marking requires the learning of an additional syntactic property.

The experimental results report that all children pass DS test, but some children fail MNC test. This shows that DS sentences are acquired earlier than MNC in spite of the lower frequency, arguing that frequency does not determine the order of the acquisition. The results are consistent with Ura’s theory of MNC, according to which MNC requires a positive setting of multiple Nom case feature parameter. This suggests that the delayed acquisition of MNC in the production data has to do with the learning of the property of multiplicity of Nom case.

Chapter 3 and Chapter 4 investigate the acquisition of scrambling. Chapter 3 reviews syntactic aspects of scrambling, and introduces three competing UG-based syntactic theories of scrambling: Grewendorf and Sabel’s (1999) Scrambling feature approach, Miyagawa’s (2003)

EPP/Focus approach, and Bošković's (2004, in press) Case-related approach. We discuss each scrambling theory from acquisitional perspectives. Chapter 4 investigates a learnability problem of scrambling. We observe that the positive evidence for scrambling in parental speech is extremely rare. Determining that input-based learning approach fails to account for how Korean children can acquire scrambling, we hypothesize that there may be a cue, which indicates the availability of scrambling, and children may have access to the cue. Considering syntactic analyses and cross-linguistic data of scrambling, we take overt Acc case marker as a possible cue. We carry out an experiment to test whether scrambling is related to the overt Acc case. We report the results which indicate a significant contingency between the acquisition of Acc and the acquisition of scrambling, supporting case-related scrambling analysis of Bošković (2004, in press). This finding indicates again that frequency did not determine the order of the acquisition. The results suggest that children must have an innate knowledge about the abstract relation between scrambling and the overt Acc case.

Further, this result makes an implication for the nature of linguistic experience that triggers learning. It has been assumed that there are so-called 'trigger data', which are sentences in the child's experience that point directly to the correct settings of the parameter and that robust amount of such data should be available to the child (Gibson and Wexler 1994, Lightfoot 1999). Our finding suggests that the child may learn some syntactic properties without hearing much of exemplary sentences, and the triggering data may not be sentences that directly exemplify the property, and can be subtle, challenging the common assumptions on trigger data.

In Chapter 5, we provide a summary of the major findings of this thesis, and discuss some implications of the findings. We discuss the nature of children's remarkable ability to match input frequency in many cases. We argue that children indeed have access to overall frequency of some linguistic elements, as argued by the statistical learning model. However, considering our acquisitional results which show that the frequency does not determine the order of acquisition and that the child must have some language specific prior knowledge to learn language particular

syntactic properties, we propose that the areas of children's success in matching input frequency are aspects of performance, not ones in which the child has to determine whether a given property is grammatical or not.

Chapter 2. Multiple Nominative Case Marking: Acquisitional Perspective

2.0. Introduction

This chapter investigates the nature of the acquisition process involved in learning of multiple Nominative Case (MNC) sentences in Korean, as illustrated in (1).

(1) a. Mary-ga cha-ga pilyohatta

M-Nom car-Nom need

‘Mary needs a car’.

b. Mary-ga Sunhee-ga chotta.

M-Nom S-Nom like

‘Mary likes Sunhee.’

c. Mary-ga cha-ga isstta.

M-Nom car-Nom be

‘Mary has a car’

d. Mary-ga ko-ga kutta.

Mary-Nom nose-Nom big

‘Mary has a big nose.’

(2) Mary-ga cha-ul sasstta.

M-Nom car-Acc bought.

‘Mary bought a car’.

The examples in (1) display a peculiar property when they are compared with the sentence in (2). Unlike typical objects marked with Acc case, as in (2), the objects in (1) appear with Nom case, together with the subject, displaying multiple instances of Nom case. ‘Stative predicates’, which are known as predicates describing states of mind, emotion or possessions, allow such an exceptional use of Nom case on the object.

From an acquisitional point of view, it may seem that children will be able to produce MNC sentences when they acquire the knowledge of two things: i) Korean has an overt Nom case marker *-ga* for the subject, and ii) a certain type of predicate requires the object to appear with Nom case, instead of Acc case. Interestingly, it has been observed from Japanese longitudinal speech data that children somehow do not produce MNC utterances even if they have learned those two things (Matsuoka 1998). We have independent evidence of productive use of Nom case as a subject marker and a considerable number of stative predicates with objects marked with Nom case. Mysteriously, however, we do not find evidence of children producing utterances of MNC in which both the subject and the object are marked with Nom case. What is the nature of the absence of MNC in the child speech?

In UG-based grammatical theories of MNC, the presence of Nom object does not necessarily grant multiple Nom case. Multiplicity of Nom case is considered as an additional property which distinguishes MNC from other types of sentences, in addition to the peculiarity of the Nom object. This predicts that children have to figure out the additional property, suggesting that the lack of MNC observed in the child speech may be due to it may be due to the learning of some abstract grammatical property of MNC such as multiplicity of Nom case. We consider two different analyses of MNC from an acquisitional perspective.

We examine distributional properties of parental input to see whether the lack of MNC in the child speech is related to input patterns, without resorting to the grammatical properties assumed in the two syntactic proposals. If the input frequency of MNC in adult speech is low as in the child speech, it may be that children simply did not produce MNC, assimilating to the adult

speech pattern. We find that MNC in the adult speech was extremely rare as in the child speech, suggesting the possibility that the lack of MNC in the child speech is not an indication of the lack of a grammatical property. In order to determine the nature of the lack of MNC in the child speech, we conduct an experiment.

The chapter is organized in the following way. In section 2.1, the case theory in Principles and Parameters model is briefly reviewed for some background on the grammatical theory of case. Section 2.2 discusses syntactic behaviors of multiple Nom case marking phenomenon. Section 2.3 reviews two competing theories of single Nom case and MNC, and further discusses their implications for acquisitional issues. Section 2.4 studies the acquisition of single Nom case. We examine both children's spontaneous speech data and their parental input patterns to consider whether the observed facts can be accounted for by an input-driven learning mechanism. Section 2.5 analyzes the developmental contour of Nom case comparing with that of Acc case. Section 2.6 investigates the acquisition of MNC. We will see that spontaneous speech data are not enough to make a conclusive argument for one approach, calling for an experimental study. Section 2.7 presents the experimental study. Finally, section 2.8 provides the conclusion of the chapter.

2.1. Case Theory

Case as a morphological object normally attached to nominal elements has been one of the major topics of study in traditional grammar. Case was regarded as a system of marking dependent nouns for the type of relationship they bear to their heads (Blake 1994). Typically, case marks the relationship of a noun to a verb at the clause level, or of a noun to a preposition,

postposition, or another noun at the phrasal level. For instance, in Korean example given in (3), *-ga/i* marks the subject, as in *John*, and *-lul/ul* marks the object, as in *Mary*.¹

- (3) John-i Mary-lul mil-ess-tta
 J-Nom M-Acc push-Pst-Decl
 ‘John pushed Mary’

However, there are languages like Chinese or Thai where overt case morphology does not appear at all. Observing the cross-linguistic differences in the richness of morphological case, Vergnaud in his personal letter to Chomsky suggests the following idea. Every lexical nominal phrase must have a particular case in order to be morphologically realized, and in order to do that they must occur in a characteristic position; Nom in subjects of a finite clause, Acc in objects of a transitive verb, etc. The particular cases have overt morphological forms in some language such as in Korean, but they may not have morphological realizations, as in Chinese.

Chomsky (1981) adopts Vergnaud’s insight, and calls the abstract requirement on nominal phrases ‘Case’.² Under this assumption, a noun phrase (NP) gets Nom Case, when it is a subject of a finite clause, and an NP gets Acc Case, when it is an object of a transitive verb. If an NP is in a position where no Case can be assigned, it is ruled out by the *Case Filter*, which demands that a phonetically overt NP has Case. For instance, consider the examples below (Lasnik 1988).

- (4) a. *It is likely John to lose the race.
 b. John_i is likely t_i to lose the race.

¹ The two forms for Nom case *-ga* and *-i* are phonologically conditioned allomorphs. *-Ga* is used when the preceding sound is a vowel, and *-i* when the preceding sound is a consonant. Acc case also has two different phonological forms: *-lul* and *-ul*. *-lul* is used after a vowel, while *-ul* is used after a consonant.

² I will be using ‘case’ to cover both the traditional ‘case’ and the abstract notion of ‘Case’, when the distinction is not crucial.

(5) a. *It was arrested Mary.

b. Mary_i was arrested t_i.

(4a) is ruled out by the Case Filter because the overt NP *John* is not in a Case position, being in a non-finite clause. The grammatical counterpart (4b) shows that *John* must move to the matrix subject position, where it can be assigned Nom case. This indicates that Case is a motivation for *John* to move the matrix subject position. The same insight is applied to account for the behavior of an object with the passivized verb. The ungrammaticality of (5a) is also explained as a Case Filter violation. The object *Mary* in (5a) is left without Case because the passive verb cannot assign Acc case to the object. This yields a Case Filter violation. The sentence becomes grammatical when the object raises to the subject position to get Case, as in (5b). This conception of Case as a motivation for NP movement provides a way to conceive of the two operations, subject raising and passivization, as a uniform phenomenon of NP movement from a Caseless position to a Case position. These facts establish the role of Case as a module to regulate the distribution of NPs in the sentence.

In the theory of Case as developed in Chomsky (1981), finiteness of clauses is encoded in a functional head called INFL, which hosts a collection of features [Tense/Agr]. A clause is finite if INFL contains [+Tense]. Finite INFL allows [+Agr], which renders the subject to agree with the verb. Nom case is licensed by this agreement. Accordingly, agreement and case are assumed to be one and the same thing. For instance, consider the following sentence.

(6) John seems to like the dog.

The nonfinite INFL in the embedded clause does not have [+Tense] feature, hence it cannot assign agreement to the embedded verb. This means that the subject *John* will not get Nom case

there, forcing it to move to the matrix subject position to get Case, where the INFL has [+Tense/Agr] to assign Nom Case.

This system of Nom case assignment works well in accounting for the facts in English, as we saw above. However, when we turn to other languages like Korean, the theory is immediately challenged.

2.2. Multiple Nom Case in Korean

Several studies of Korean grammar have questioned whether INFL in Korean is indeed a functional category which can license Nom case. Kang (1986), Kim (1991) and others proposed that Korean INFL lacks relevant features to license case and agreement. This line of proposal is based on the claim that neither functional feature [+Agr] nor [+Tense] is a licenser of Nom case or agreement in Korean. Korean has two candidates for agreement morphology: plural agreement *-tu* and honorific agreement *-si*. However, Kang (1986) argues that these two agreement morphologies do not indicate the same status of subject-verb agreement as in English because they are optional, unlike English in which the lack of agreement induces ungrammaticality. The status of [+Tense] feature has also been doubtful. Kim (1991) claims that Tense is only a place holder of tense morpheme, rather than a syntactic category capable of assigning Nom case, providing several cases in which finiteness/non-finiteness does not seem to correlate with the presence of Nom case.

Apart from the absence of ‘real’ agreement morphologies and the apparent presence of Nom case in non-finite contexts, there are two other notorious phenomena are often used as a strong argument for the lack of INFL as a functional category in Korean. They are multiple Nom case property and dative subject sentences, as given below.

(7) a. Mary-ga cha-ga pyilyohatta. (Multiple Nom Case)

M-Nom car-Nom need

‘Mary needs a car.’

b. Mary-ekey cha-ga pyilyohatta (Dative Subject)

M-Dat car-Nom need

‘Mary needs a car.’

Both constructions exhibit some peculiar case patterns. In MNC sentence (7a), the object as well as the subject is marked with Nom case, forming multiple Nom case. In DS sentence (7b), the case of the object remains the same, but the subject appears with Dative case instead of Nom case.

Interestingly, the two types of sentences share some properties. Both types are restricted to predicates which describe the status of mind, emotion, and possessions: *mwupseptta* (fear), *chotta* (like), *pilyohatta* (need), *itta* (have). etc. These predicates are traditionally defined as ‘stative predicates’. This peculiar property of the predicates determines the peculiar nature of the semantic role of the subject, imposing that the predicates do not have capacity to assign an agent role. Hence, the subject with those predicates ends up getting a role which undergoes some experience. This role is often known as *Experiencer*.

There are other instances of MNC which do not allow alternation with Dative case on the subject. Those are MNCs with inalienable possession relations, as shown in (8a).

(8) a. Mary-ga ko-ga kiltta.

M-Nom nose-Nom long

‘Mary has a long nose.’

b. *Mary-ga cha-ga bissatta.

M-Nom car-Nom expensive

‘Mary has an expensive car.’

In (8a), two Nom marked element, *Mary* and *ko* (nose), are in an inalienable relation. When *nose* is replaced by an object that is alienable from *Mary* such as *cha* (a car), multiple Nom is no longer possible, as shown in (8b).

Dative subjects are found in many languages ranging over various language families: Germanic languages (Icelandic, Faroese, German, Old English), Romance languages (Spanish, Italian), South Asian languages (Malayalam), Indo-Aryan languages (Hindi, Nepali, Maithili), Slavic languages (Russian, Polish, Serbo-Croatian, and Bulgarian), and Altaic languages (Turkish, Japanese, Korean). The fact that among those languages with dative subjects, only Korean and Japanese exhibit multiple Nom case marking suggests that Multiple Nom Case constructions forms a subset of these languages.

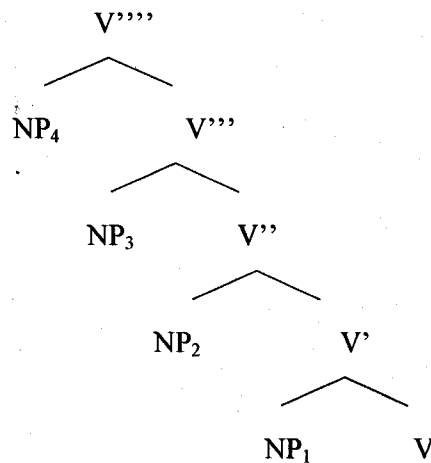
To summarize, under the standard case theory, the case patterns above display three interesting properties. First, the subject can be marked with dative instead of Nom case. Second, the object can be marked with Nom case instead of Acc case. Third, more than one argument in a single clause can receive Nom case. These facts together with the argument concerning the lack of agreement feature lead to a proposal, which hypothesizes that Korean INFL does not have Nom assigning feature. Under this radical proposal, languages vary as to whether INFL has Nom assigning feature or not. The Korean INFL is parameterized to the negative setting. Nominal elements get assigned Nom case marker by default, independently of INFL. We call this proposal *Default Case approach*. A competing hypothesis is proposed, with the assumption that the Korean INFL has Nom case assigning feature just as in English, but an additional property of INFL such as the possibility of multiple feature checking is parameterized to give rise to the apparent peculiar case patterns. We call this *Multiple feature checking approach*. In the following sections, we review these two competing syntactic theories and their implications for learnability.

2.3. Two Competing Views on Multiple Nom case

2.3.1 Default Nom Approach

Kang (1986) proposes an analysis of Korean case patterns parameterizing the basic phrase structure of language. He argues that languages differ in their choice of the head of a sentence. In a language like Korean, S is the projection of V, instead of INFL. The head V, being a lexical category, is assumed to allow an unbounded number of arguments in specifier positions, unlike the functional head INFL, which requires only one specifier. This is schematized below.

(9) Kang (1986)



Based on the above assumption of parameterized phrase structure, Kang proposes the following case marking rules to account for the Nom case patterns in Korean.

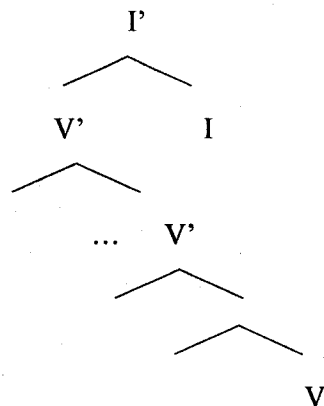
- (10) a. An NP which is a sister of [-stative] V is assigned Acc in the course of derivation from D-structure to S-structure
- b. Nom case is assigned to all non-Case marked NPs.

According to (10a), if a verb is [+stative], the NP1, which is a sister of V in the structure (9), the verb does not get assigned Acc case. The rule (10b) applies by default, and assigns Nom to the

NP1 and all other non-case marked NPs (NP2, NP3 and NP4). This gives rise to instances of MNC. Notice here that ordering of the rule application matters. The rule (10a) has to apply before the rule (10b). Parameterization of the head of S to V grants multiple specifiers, and the default case marking rule (10b) distributes Nom case to all those specifiers.

Fukui (1986, 1988) makes a similar proposal for Japanese, except that he adopts the hypothesis that Japanese has a functional category INFL, but only as a place holder of tense morpheme. He argues that the INFL in Japanese is “very defective” and contains no agreement features. On his assumptions the presence or absence of agreement feature in a functional category has critical implications. Only a functional head with agreement features has the ability of taking a specifier, closing off the projection. Further, one and only one specifier is allowed because the agreement between INFL and the specifier is strictly one-to-one relation due to what he calls *uniqueness* principle. Under this theory, the Japanese INFL will never license specifiers, lacking agreement features. Instead, Japanese can allow reiteration of NPs at the V' level because lexical categories never close off. This drives the following phrase structure.

(11) Fukui (1988)



Nom case is assigned to all specifiers which are the sisters of the v', which can be reiterated. This proposal is in the same vein as Kang (1986), apart from the recognition of the category of INFL.

Fukui and Takano (1998) reanalyze the intuitions of this approach with the terms in Minimalist syntactic theory.³ In the Minimalist theory, case feature is to be eliminated before an operation called Spell-out, a point of splitting the syntactic derivation to two interface levels (phonetic form and logical form), because it is not interpretable in either level. Case assignment is replaced by case checking operation, according to which case features in a nominal element and a functional element are matched, and then eliminated under a certain configuration. Fukui and Takano propose that UG provides two different ways of eliminating Case features: Checking and Spell-Out. Checking is employed by abstract Case. The checking takes place within the domain of a functional head. For instance, finite T provides checking domain for Nom checking and the light verb *v* for Acc checking. Spell-out option is reserved only for morphological case. Assuming that morphological case, having phonetic content, is visible to the Spell-Out operation, they suggest that this enables the morphological case to be eliminated by the Spell-out operation. This option is not possible for abstract Case. The abstract case has to be eliminated before Spell-Out, because it does not have phonological features that can be visible to the Spell-Out operation. In a word, morphological case is eliminated simply by spelling-out the case, while abstract Case must establish a checking relation with an appropriate functional head.

Accordingly, Nom case particle and Acc case particle in the nominal phrase are checked by Spell-out operation naturally. What about the case assigning features in a functional head such as T?⁴ Following Kuroda (1983), Fukui and Takano assume that Japanese Nom case is not

³ Fukui and Takano (1998) share some insights of Kuroda (1986, 1988) in their theory. Kuroda proposes that languages differ as to whether agreement is forced or not, which he calls [+/-Forced Agreement] Parameter. In his theory, agreement must be in a one-to-one relation, restricting one functional head to agree with one maximal projection. English, being a +Forced agreement language, does not allow more than one element to occur with the same case, since a functional head is forced to undergo agreement and the agreement is a relation between a single head and a single functional projection. On the other hand, Korean or Japanese are not subject to this constraint, because agreement is not forced in these languages. Languages with the negative setting of the Agreement Parameter are free from the one-to-one agreement constraint, and have linear case marking available, which allows multiple case.

⁴ As for the Acc case in V, Fukui and Takano (1998) assume that Acc case in Japanese cannot undergo checking because Japanese lacks *v* to V movement, unlike English. Consequently, they suggest that Acc case in Japanese is an inherent case, which does not require case elimination.

contingent on tense, which they relate to the lack of subject-verb agreement. Given this assumption, they arrive at the conclusion that T in Japanese does not have Nom case assigning feature. This theory takes the presence or absence of morphological case in a given language as a determining factor to decide on which option of case elimination is available. The defective status of INFL is matched with having Nom morphological case. This analysis is slightly different from the previous theories in that the previous versions of default case mechanism seem to build the theory on the defective status of INFL, while this analysis focuses on the availability of Nom case particle in a given language.⁵ This makes the learnability easier because the child just needs to recognize the Nom case particle to figure the status of INFL.

Now, let us see how this theory accounts for the MNC property. Fukui and Takano (1998) derive MNC as a consequence of having Spell-Out option for case elimination, which becomes available with morphological Nom case. Case elimination by Spell-Out can apply to any number of NPs with morphological cases since it does not form checking relation with a functional head. As a result, multiple occurrences of morphological Nom case are possible. On the other hand, MNC is not possible with abstract Nom Case. In the checking operation it is impossible for T to check more than one Nom noun phrase because feature elimination by checking takes place as soon as the proper configuration is formed. The Nom feature of T is eliminated when one noun phrase is inserted in the checking domain of T, and becomes unavailable for further checking. This explains why English does not allow MNC, for instance.

2.3.2 Multiple Feature Checking Approach

Ura (1996, 2000) make an opposing proposal to the Default case approach. His theory is based on what he calls 'grammatical function splitting'. He makes a distinction between grammatical functions and grammatical relations. Grammatical functions (GF) are such properties as the ability to control PRO, the ability to bind reflexive, etc. Grammatical relations

⁵ I thank Diane Lillo-Martin for this observation.

(GR) are the terms traditionally defined as subject and object. It has been standardly assumed that if an element A counts as having the GR subject, then A is expected to have the set of GF that are linked with the GR subject. Ura recognizes several phenomena which challenge this assumption. He shows that there are cases in which grammatical functions of a single head are split into two different elements. One such case is instances of Dative subjects. The NP with Dative case can bind subject-oriented anaphor, as in (12a), and also has the function of controlling PRO, as in (12b). On the other hand, in languages with subject-verb agreement such as Tamil, the ability of inducing subject agreement on the finite verb is borne by the Nom marked object, as in (13).

(12) a. John-ekey_i Harry-ka casin-uy_i sengkong-ul wihayse philyoha-ta Korean

J-Dat H-Nom self-Gen success-Acc for need-Decl

‘John_i needs Harry for self’s_i success.’ (O’Grady 1991:102)

b. PRO_i haksang-i-myense, John-ekey_i manun ton-i pyilyoha-ta Korean

student-be-though John-Dat much money-Nom need

‘Although PRO_i being student, John_i needs much money.’ (O’Grady 1991:103)

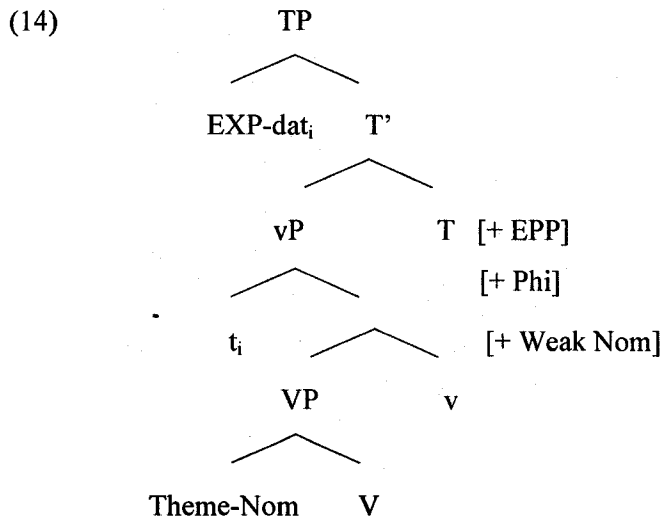
(13) Kumaar-ukku irantu paiyan-kal iru-kkir-aarkal Tamil

K-Dat two boy-PL-Nom be-pres-3PL.EP

‘Kumar has two boys.’ (Lehmann 1993:189)

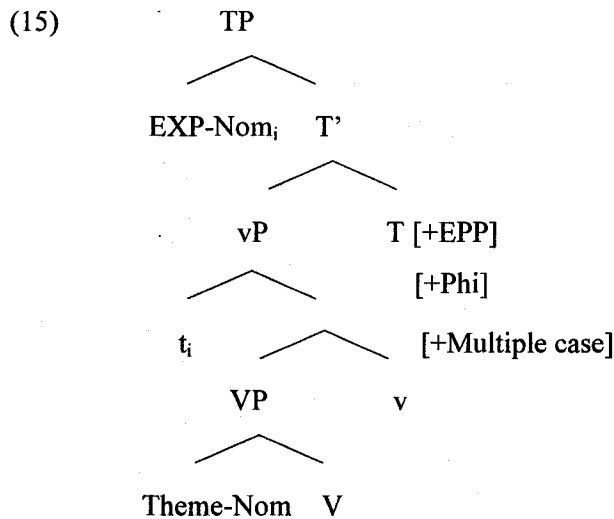
Ura calls this a grammatical function splitting phenomenon. Observing this kind of phenomenon, Ura determines that we can no longer relate grammatical relations to grammatical functions in a deterministic fashion. He proposes that the checking theory of Minimalist theory (Chomsky 1995) provides a useful tool to account for the grammatical function splitting phenomenon because under the checking theory, structural relations are defined by feature checking relations, not by a fixed structural position. This means that different kinds of feature checking relation can yield

different grammatical functions. Let us see how this theory applies to DS first before we take a look at the analysis of MNC.



Ura assumes that the experiencer (EXP) is generated at the spec of a light verb which takes a VP with theme in its complement position. The light verb here is incapable of assigning Acc to a Theme. EXP is assigned dative case as its inherent case by the psych-predicate. A strong EPP feature of T attracts the D-feature of the EXP for EPP feature checking. Consequently, the EXP is raised to the spec of T. Phi-feature of T is checked against the Experiencer. Weak Nom case feature of T checks off the Case feature of the Theme at LF. EPP-feature and Phi-feature are checked against the EXP, while the case feature of T is checked against the Theme.

This splitting of feature checking relations is extended to the derivation of multiple Nom case construction. However, MNC requires an additional property to allow the multiple Nom instance. Ura proposes that a given functional head can have the option of allowing multiple features. He suggests that languages are parameterized as to whether a functional head can have this possibility. If a given functional head has a positive setting of the parameter, the head can enter into multiple feature checking relation, allowing multiple specifiers. Let us see how Ura accounts for MNC using the parameter.



In this case, the predicate does not assign the inherent dative case to the EXP. The EXP is raised to spec of T and checks EPP-feature of T and the phi-feature of T, essentially the same way as in DS. Now, the EXP must check its Nom case feature. The weak Nom case feature in T checks off the Nom case feature of the EXP and the Nom case feature of the Theme. The property of [+multiple Case] makes it possible.

2.3.3. Implications for Learnability

The two syntactic analyses reviewed above differ in what aspect of grammar can be parameterized. Ura's (1996, 2000) Multiple feature checking approach takes structural case licensing to be a universal principle of grammar, and assumes that a functional head T is parameterized as to whether T allows multiple feature checking or not. This Multiple feature checking parameter intends to loosen the one-to-one checking relation, allowing one-to-many checking relation. This makes MNC available in a language. On the other hand, Fukui and Takano's (1998) Default case approach takes one-to-one checking relation to be a universal property of grammar, and allows a parameterization as to whether a functional head T has an ability to check structural case at all. In other words, the functional head T has an option with respect to checking itself; T may or may not have a Nom case feature. When T comes with the

Nom case feature, it forms a checking relation with an appropriate nominal phrase. However, when T lacks the Nom case feature, checking does not take place because it does not have formal feature to trigger case checking. Consequently, T does not form case checking relation with the nominal phrase. Assuming that in Japanese or Korean, Nom case is not contingent on tense, unlike English Nom case, Fukui and Takano argue that T in Korean and Japanese does not have Nom case assigning feature. Nom case particle on the nominal phrase is eliminated by Spell-out operation. In this approach, the existence of MNC follows from the nature of the Spell-out operation, which is free from the restriction of one-to-one checking relation.

Both syntactic analyses achieve the same empirical coverage for the relevant data even if they diverge in a fundamental assumption on what aspects of grammar can be parameterized. How do we tease the two approaches apart? A more in depth syntactic analysis may resolve this situation, but here we choose to turn to learnability question to evaluate the two analyses; we ask whether both theories are equally learnable.

The two approaches make different acquisitional predictions, even if they both assume that multiplicity of Nom is a unique syntactic property. Under Ura's Multiple feature checking approach, T has Nom feature as a general property in all languages, and has a parametric option as to the value of multiple feature checking. It implies that the relevant positive evidence for the multiple feature checking parameter would be the instances of MNC. If we assume that negative evidence is not available, we expect that the child starts with negative setting of multiple feature checking parameter, and later changes to the positive setting, on the basis of MNC utterances. This predicts that there will be a stage in which the child will not produce MNC because they have not set the positive value of the parameter yet.

In the Default case approach by Fukui and Takano, languages differ as to whether Nom case requires structural licensing or not. This parameterization is related to whether a functional head T has Nom case assigning feature not. In a language in which Nom case does not require structural case licensing, Nom case particle is available. This allows Spell-out option of case

feature elimination, which is free from the restriction of one-to-one checking relation. The Spell-out option makes MNC available as a free gift. This predicts that Nom case particle is enough a positive evidence for MNC at the same time being a positive evidence for the availability of the Spell-out option. In other words, when the child learns that Nom case particle is not contingent on Tense, which is evidenced by lack of subject-verb agreement, they will know that MNC is also available. As long as the child learns Nom case marker and the relevant predicates along with the requirement of Nom case marking of the object, they should know that MNC is possible in their grammar.

2.4. Acquisition of Single Nom Case

In this section, we review previous studies on the acquisition of single Nom Case. We analyze acquisitional facts such as the order of acquisition and the developmental contour with respect to Acc case to determine whether those facts can be explained away by input-driven learning approach. Particularly, we compare the child speech with parental input to determine whether input frequency patterns can account for the facts. We will find a surprising fact that the child speech exhibits a close tie with the input frequency.

First, we review Matsuoka's (1998) findings on the acquisition of case markers with Japanese learning children, and compare them with several acquisition studies on Korean case markers. Sources of the Korean data we rely on are Kim (1985), Chung (1994), and Cho (1982). Besides the reports from the previous studies, we also examine Jiwon's corpus, which recently became available in CHILDES system. Jiwon's corpus is of a small size, which consists of only 5 files recorded over about one month period between 2;0,13 and 2;1,10 with a weekly interval. This corpus does not contain much of the child speech, but provides us with valuable information about properties of parental language input.

Matsuoka (1998) investigates acquisition of case markers with Japanese speaking children to determine whether the acquisition of case markers in Japanese is input-driven or guided by some innate knowledge of language. In her study, she finds two important facts on acquisition of case markers. First is the consistency in the order of the acquisition among three Japanese children. Matsuoka reports that for all three children she observed, Nom case appears earlier than other case markers such as Acc or Dative case. This is shown below.

(16) Table 1. Age of the first occurrence of case markers (Japanese)
(Matsuoka 1998:70)

	AKI	KAN	Sumihare
Nom (<i>ga</i>)	2;2;22	2;2;3	1;11
Acc (<i>o</i>)	2;9;7	2;2;7	2;1
Dat (<i>ni</i>)	2;4;19	2;2;14	2;1

Importantly, she observed that the order of acquisition of case particles was uniform regardless of the amount of input concerning the use of case particles. The table below shows that two children (AKI and KAN) received more occurrence of Dative case than Nom case in the input; nonetheless, they both followed the general pattern of acquiring Nom case before Dative case. This is mysterious if frequency is the main factor to shape the developmental course of grammar.

(17) Table 2. Percentage of the three particles in adult speech (Japanese)
(Matsuoka 1998:69)

	AKI	KAN	Sumihare	All corpora
Nom (<i>ga</i>)	44%	31%	38%	40%
Acc (<i>o</i>)	6%	15%	32%	12%
Dat (<i>ni</i>)	50%	54%	30%	48%

What can explain the early acquisition of Nom case if children are not relying on frequency? Matsuoka (1998) investigates this question, examining the timing of the emergence of Nom case marker. Assuming the checking theory of Case in the Minimalist syntactic theory in Chomsky (1995), according to which Nom Case is licensed by finite Tense feature, Matsuoka hypothesizes that the emergence of Nom case should be contingent on Tense morpheme. She predicts that Nom case will not be available until the licensing element Tense is fully productive. The following table shows her data regarding this prediction.

(18) Table 3. Usage of Tense morpheme by children (Japanese)

(Matsuoka 1998:80)

	AKI	KAN	Sumihare
Nonpast	1;10	2;2	1;11
Past	2;1	2;2	1;11
Nom	2;2	2;2	1;11

Two children (KAN and Sumihare) were already producing both Tense morpheme and Nom case in their first files, making it impossible to determine their order of appearance. However, the data from AKI shows non-past/past ending appeared earlier than the Nom case. Matsuoka reports that the result was statistically significant at the .001 level according to sign test.⁶ Based on this result, she suggests that the abstract checking relation between Tense morpheme and Nom case as a part of UG guides Japanese children's acquisition of Nom case markers, arguing for the view that Nom case is licensed by Tense feature (Takezawa 1987, Ura 1996, 2000).

Now, we consider Korean data to check whether the two main findings in Matsuoka's (1998) study are also attested in Korean children's speech. First, concerning the order of case

⁶ The sign test was run to determine the probability that the present tense ending appears before the first use of Nom case particle by chance (See Matsuoka 1998 for the detailed figures).

markers, all three studies we reviewed show that Nom case is acquired significantly earlier than any other case markers, confirming Matsuoka's finding. Kim (1985) reports that all five children in her study began to supply Nom very early between 1;8 and 2;0, which was around the onset of two word utterances, and no case marker appeared at the one word stage. Cho (1982) reports a similar fact. Cho's children first produced the Nom marker as early as 1;7 and as late as 1;11. The appearance of Acc case comes four or five months after the onset of Nom case. The specific timing of the first occurrence of Dative was not available.

In order to check any possible interaction with the input patterns, we checked the proportions of these case markers in the input using Jiwon's corpus. Interestingly, Nom was far more frequent than both Acc and Dative.

(19) Table 4. Percentage of the three particles in adult speech

	Jiwon (2;0)
Nom (<i>ka/i</i>)	390 (71%)
Acc (<i>ul</i>)	77(14%)
Dat (<i>ey</i>)	80 (15%)

Notice here that the proportion of Dative case is much smaller than that of Nom case, unlike the data in Japanese. Jiwon first produced Nom case, and none of the other case markers appear in all 5 files, except the topic/focus marker *-nun*, which was as frequent as Nom case in the input.⁷ The early appearance of the topic/focus marker seemed to be related to frequency.

These Korean data seem to suggest that the acquisition order of case markers is determined by input frequency. However, we cannot extend this to Japanese because we saw

⁷ The frequency of Nom case and Topic/focus marker *-nun* are shown below.
 Nom 390
 Topic/Focus 400

from Matsuoka's data that the frequency of Dative case was higher than Nom for two children, but both children acquired Nom earlier than Dative case. Based on the Japanese facts, we may argue that frequency does not account for the facts, and take this as a case which shows the importance of comparative acquisition study. However, without investigating the status of Dative case in both languages, we cannot draw any conclusive argument at this point.

Instead, we turn to the issue of concurrent emergence of Tense morpheme and Nom case. Kim (1985) reports that one Korean speaking child began to produce and productively use Nom at 1;9,28, and the same child acquired past tense marker at 1;9. Another child who began to produce and productively use Nom at 2;0 is reported to have acquired past tense at 1;11. This result seems fairly consistent with the observation made by Matsuoka (1998). Jiwon's corpus does not show the stage of productive use of Nom or past tense marker yet, but the first use of Nom and past tense morpheme appeared in the same file.

The results on the concurrent emergence of Nom case and Tense morpheme in both languages seem to constitute a strong argument for the UG-based learning approach. Given that tense morpheme and Nom case are two unrelated grammatical elements, the contingency in their acquisition is not likely to occur by chance. This means that the child must have the prior knowledge of the abstract relation between the two elements.

Can we then conclude from the above fact that tense morpheme is indeed the mediator responsible for the early appearance of Nom case? In order to draw this conclusion, we need to rule out the possibility that input frequency of tense morpheme may have affected the result. In order to check this possibility, we examined the frequency of the use of past tense in Jiwon's corpus, and compared it with the timing of the first occurrence of the Nom particle.

(20) Table 5. The frequency of Nom and Past tense in the input

	Jiwon (2;0)
Nom (<i>ka/i</i>)	390
Past tense (<i>-ess</i>)	403

Consider the data shown in the table above. Surprisingly, the past tense marker was slightly more frequent than Nom case. Observing this, one is very likely to argue that the close tie in the emergence of Nom case and Tense morpheme is actually due to their compatible frequencies in the input rather than any abstract licensing relation between Tense and Nom. Since we cannot exclude the possibility that the concurrent emergence of Nom and Tense is due to the compatible frequency, the issue of whether Nom case in Korean and Japanese is structurally licensed or not, as argued in Matsuoka (1998), cannot be determined. In the following section, we turn to examining the developmental contours of these elements to see how far this input-driven account can go.

2.5. Developmental Contour of Nom case

In the previous section we focused on the first occurrence of Nom case particle. In this section, we take a close look at the developmental contours of Nom case compared with Acc case marker. But before we discuss the developmental data, we need to point out one important difference between Nom and Acc case, which we believe to be crucial to understanding the developmental facts. Acc case is freely droppable when the object appears in the canonical order, while Nom marker is always required unless other operation such as topicalization or focalization is applied.⁸ The following data demonstrate this point.

⁸ When an element is topicalized or focalized, the overt marker is attached to the element, replacing the case markers.

(21) Kuno (1973) and Saito (1985:208)

a. John kita no?

J came Q 'Did John come?'

b. dare-ga kita no?

who-Nom came Q 'Who came?'

c. *dare-wa kita no

who-Top came Q

d. *dare kita no?

who came Q

The contrast between (21b) and (21d) shows that the ungrammaticality of (21d) is due to lack of Nom case. This means that the bare subject in (21a) is not an instance of Nom drop, but topic marker drop. Unlike Nom, Acc case can be omitted when the object is in the canonical position, as shown in (22).

(22) a. John-i cha-(lul) sasassta.

John-Nom car-Acc bought.

'John bought a car.'

b. John-un cha-(ul) palassta.

J-Top car-Acc sold

'John sold his car/cars/a car/'

The data in (22) also show that the presence or absence of Acc case is not related to any semantic effect such as definiteness or specificity. The presence of Acc case is not conditioned by any

linguistic factor, unlike in Turkish.⁹ If the Acc case is optional, what is the rate of use of the case markers? The following table shows that only 20-30% of objects in the parental input appear with Acc case.

(23) Input frequency of Case markers (Chung 1994, Kim 2002, Clancy 1985)

Table 6. Rate of Nom/Acc out of potential contexts in adult-to child speech¹⁰

	Korean	Japanese ¹¹
Nom	80%	70%
Acc	20-30%	20%

For the remaining 70% of objects, the child hears them without Acc marker. This means that children receive reliable evidence for grammatical use of Nom case, but as for Acc case, they are exposed to highly inconsistent input.¹²

The inconsistent input of Acc raises a question: how do Korean children acquire Acc case in such inconsistent input? We can think of at least three logical possibilities. One possibility is that given that the child hears a significantly greater proportion of objects without Acc case, they may conclude that the object without Acc is the grammatical form, and the object with Acc case is not fully grammatical. A second possibility is that the child may hypothesize that Acc marker is the grammatical option, and determine that somehow the adult is not being consistent.

⁹ It is known that in Turkish the presence of Acc case has the effect of specificity interpretation (Kornfilt 1994).

¹⁰ Chung (1994:102) reports the rates of Nom and Acc case in news broadcasting speech. The Nom case was produced 100% of the time, and the Acc case, 96% of the time.

¹¹ This rate is from Miyazaki (1979).

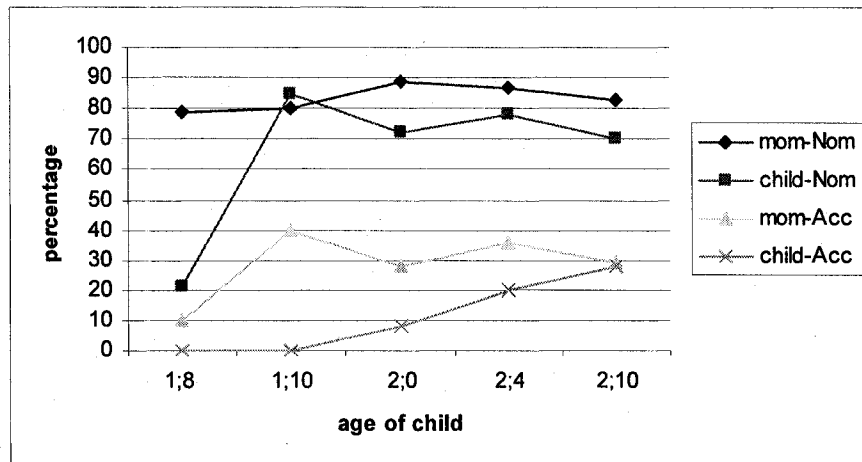
¹² Note here that the lack of Acc case does not create a problem with interpretation, because if the child knows the basic word order, which is very likely, they will be able to comprehend the meaning of the sentence.

Alternatively, the child may know that both forms of the objects are equally grammatical, and somehow this prior information prepares them to match the adult-level of frequency.

Now, let us see the actual developmental data. All studies we reviewed report that for both Nom and Acc markers, children began with a very low frequency compared to the frequency used in adult speech, but later they reached adult-level of frequency. The following two graphs show the developmental contour of each case marker by two children compared with their adult input.

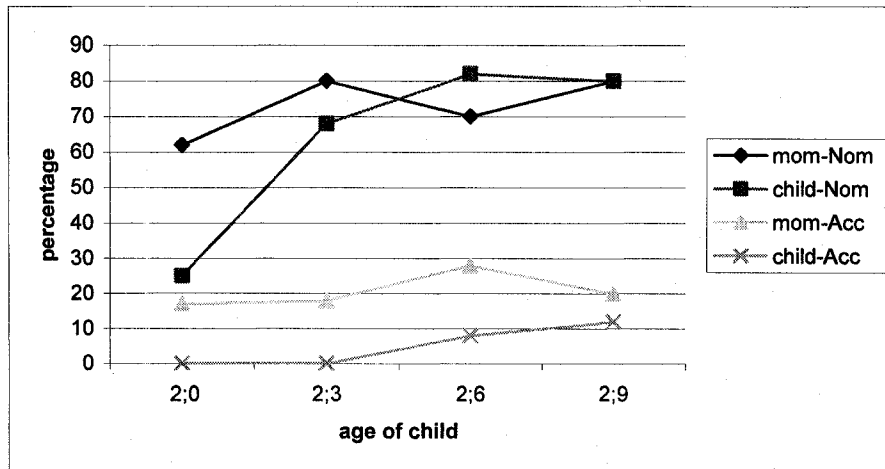
(24) Figure 1. Percentage use of the case marker by Hyuck and his mother

(Chung 1994: 99)



(25) Figure 2. Percentage use of the case marker by MJ and his mother

(Chung 1994: 99)



The two graphs show similar patterns in the way the child reaches the adult level frequency of the two case markers. Acc case seems to undergo a gradual change from non-adult frequency to adult-level frequency, while Nom case shows a steep increase. The contrast between Nom case and Acc case clearly shows that gradual learning is not a general property of grammar development.¹³

We want to take up two questions concerning the above developmental facts. The first issue is the difference between Nom case and Acc case in the way children arrives at the adult-level of frequency. Why is it that the development of Nom case shows a relatively rapid change, while the development of Acc shows a gradual increase? The second issue concerns the remarkable success of the child in matching the adult-rate of Acc case, which is only about between 20-30%.

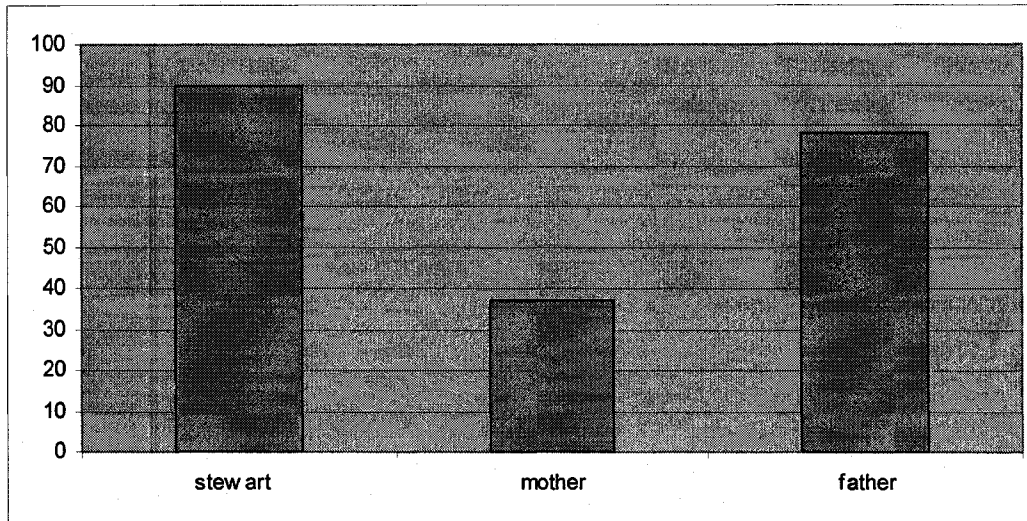
¹³ When we look at developmental patterns in other aspects of grammar, we seem to find a mixture of both types of change. It is known that children learning non-pro-drop languages start with omitting overt subjects, and gradually supply more overt subjects (Boster 1997, Hyams 1981 and Yang 2003). However, language particular properties such as wh-movement or basic word order show up early and do not seem to undergo gradual change.

Of the three possible developmental patterns considered earlier, what Korean children seem to entertain is the last possibility. They seem to take the form used with the lower frequency as well as the form used in the higher frequency as equally grammatical. Besides that, they even end up matching the frequency of adult usage with such an exact precision. How can we account for this fact? What allows them to determine that in spite of a low 20-30% of occurrence, the object with Acc marker is fully grammatical, and to guide them to match the adult-level frequency? Can we explain this fact with an assumption that all that children need is to find one major pattern, and match it with the frequency displayed in the input?

An interesting contrast arises when we compare this case with another case of input inconsistency. Ross and Newport (1996) study the acquisition of ASL verbs of motion by deaf children who were exposed to inconsistent use of morphemes on motion verbs. The parents were hearing parents who have learned to sign only with their child, and their fluency in ASL was often extremely limited. All of the parents use the morphemes to some degree, but vary in the consistency with which they use the morphemes in their required contexts. They either omit the required morphemes, or replace them with ungrammatical forms. However, the comparison of the children's performance of this morphology with the performance of their parents presented a strikingly different picture. Newport and Aslin (2002) provide the following data on the use of the morpheme for two of these children and their parents.

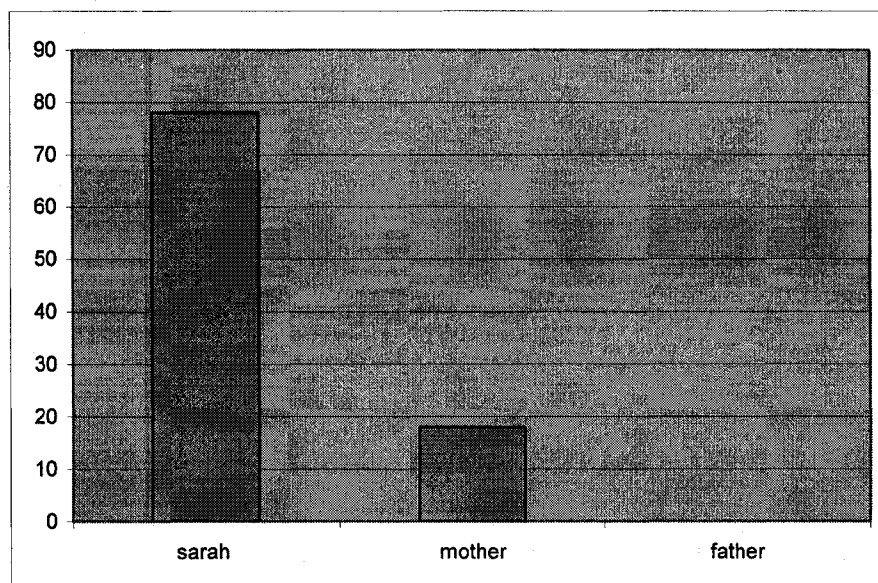
(26) Figure 3. Percent correct production of ASL movement morphemes, for Stewart as compared with his hearing parents.

(Newport and Aslin 2002:14)



(27) Figure 4. Percent correct production of ASL movement morphemes, for Sarah as compared with her hearing parents.

(Newport and Aslin 2002:14)



As shown in the figures above, the two children show near ceiling performance, in spite of the fact that their parents show quite a variable level of performance on these morphemes. Especially, Sarah's performance is surprising given her parents' poor performance.

Observing this 'creolization effect', Newport and Aslin consider two hypotheses. One is that the child knows innately that natural language morphology is deterministic, and acquires this morphology in accord with this knowledge. The other hypothesis is the assumption that the child may have a tendency to sharpen, and regularize distributions that they find from the input, as they learn. They advocate the latter hypothesis (but without much convincing evidence), suggesting that the learner may be capable of picking up the major form in the distributions, and tend to regularize it. They argue that this is how the child forms rules out of statistical data, suggesting that children exposed to probabilistically organized input use these statistical distributions to acquire 'rules'.

This phenomenon together with Korean children's acquisition of Acc marker makes an interesting contrast. We saw earlier that when the Korean child hits on the adult level of frequency, they remain there. This fact alone may be explained by input-driven learning. One could assume that all that the child needs is to find a pattern and to match it up to the frequency displayed in the input. However, this picture does not carry over to the case of deaf children's acquisition of ASL verbs of motion, because the deaf children regularized the inconsistent verbal morphology beyond the level of input frequency. Regularization mechanism may work for this case, but it raises a question for Korean Acc: if the child has a tendency to regularize a pattern, why don't Korean children regularize Acc marker, increasing the frequency up to 90%? This situation tells us that regularization and input-matching both get us into trouble. How can we account for this apparently conflicting situation?

The only way we can make sense of it is to assume that the child has a prior knowledge about what is obligatory and what could be optional. The child may already know that the verbal morphology is obligatory, and this prior knowledge tells them to supply the morphemes in all

obligatory contexts, despite inconsistent input. Likewise, the child may know that Acc marker can be optional. This knowledge guides them to stay in the adult level of frequency. If they have prior knowledge that Acc case marker drop is one of the possible options, they would not go on regularizing since the object without the overt Acc case is already legitimate according to their internal grammar.

Now, how do we account for the difference between Nom and Acc in their developmental pattern? We speculate that it is related to whether a case is obligatory or optional. If a case is obligatory, the child can produce it in obligatory contexts quickly after they learn the case form.¹⁴ They do not need to consider the adult frequency. However, if a case is purely optional, the situation is different. The child will have to pay attention to the adult frequency, and this seems to make them slow down in the development, resulting in the gradual change. This means that the child will have to come with a prior knowledge about which case is obligatory or which case is optional.

To summarize, developmental facts on two case markers seem to present a mixed picture. We saw that the order of acquisition of case markers shows a close tie to the input frequency. The case marker which is acquired early happens to be also very frequent in the input, suggesting the possibility that we can dispense with the need for UG-based learning at least for the acquisition of morphological cases. However, a close examination of the developmental contour of case markers presents a puzzle to the purely input-based learning model. Children begin to produce case markers with a low frequency, but gradually increase the frequency up to the adult level. Even though this fact seemed to be evidence for input-matching learning at first sight, it presents a problem when compared with Ross and Newport's (1996) study on deaf children's learning of ASL verbs of motion. They reported that some deaf children who were exposed to extremely

¹⁴ This does not mean that the child would produce Nom case in all overt subjects. They may use the subject without any grammatical marker. Such case can be considered to be an instance of topicalization, because topic marker can be dropped. We assume that this contributes to the less than 100 percentage of Nom case rate in the adult speech.

inconsistent input, go beyond the adult level of frequency, regularizing the inconsistent verbal morphology provided in the input. These facts suggest that the child must have a prior knowledge about optionality for a given grammatical property in order to distinguish between when to ‘regularize’ and when not to.

2.6. Acquisition of Multiple Nom Case Marking

The discussion of the acquisition of single Nom case left it unresolved whether Nom case in Korean and Japanese requires structural licensing or not. In this section we investigate the acquisition of MNC. The findings on MNC will indirectly provide an answer to the question since the two different theories of MNC diverge on that very issue.

MNC involves two peculiar facts. First, it appears in stative predicates. Second, the object is marked with Nom case, not Acc case. Given these two facts, one may assume a simple acquisition scenario of MNC. The child needs to learn two things: i) Nom case *ga* marks the subject, and ii) Predicates which belong to stative predicates require Nom case for the object, not Acc case. We first review Matsuoka’s acquisition study on MNC in Japanese to see whether this simple acquisition scenario works.

2.6.1. Matsuoka’s (1998) Acquisition Study on Japanese MNC

Matsuoka’s (1998) study on acquisition of case particles by Japanese children reviewed in the previous sections also investigated the issue of the acquisition of MNC. Matsuoka searched for potential MNC utterances with a particular type of stative predicates known as psych-predicates: *wakaru* (to understand), *iru* (to need), *dekiru* (to be able to handle), *hoshii* (desirable), *suki* (to be fond of). She observes that psych-predicates which require Nom object appear about one or two months after the first use of Nom. The following table provides the age of the first appearance of psych-predicates along with the first appearance of Nom objects.

(28) Table 7. The first appearance of stative predicates (psych-predicates) and Nom objects

(Matsuoka 1998:84)

	Stative predicates	Nom Objects
AKI	2;4;29	2;6;15
KAN	2;3;14	2;4;25
Sumihare	1;11	2;3

Interestingly, she did not find any utterances in which both the subject and the object were marked with Nom case. She reports that there were no MNC observed in any corpus. As indicated in the table below, utterances with stative predicates appear with only a single Nom case, mostly marked on the objects.

(29) Table 8. Multiple Nom Case constructions in child speech

(Matsuoka 1998:85)

	Stative (psych-) predicates	Nom object	Multiple Nom
AKI (1;5-3;0)	199	18	0
KAN (2;2-3;0)	42	3	0
Sumihare (1;11-3;3)	48	17	0

This is mysterious from a simple acquisition scenario that the child will be able to produce MNC when they know that Nom case marks the subject and stative predicates mark the object Nom, not Acc. Matsuoka interprets that the lack of MNC in the child speech indicates their lack of grammatical knowledge of MNC. Specifically, she assumes that the data reflect children's grammar in the stage where children have not learned multiplicity of Nom case yet. Following Ura's (1996) theory of multiple feature checking parameter, she argues that the children have not set the parameter to the positive value; in other words, in their grammar, the Tense head carries

only one set of formal features to check Nom case, licensing Nom case on the subject or the object, but crucially not on both elements at the same time.

She mentioned that not many stative predicates appear with two arguments. There were 19 for AKI, 3 for Kan, and 18 for Sumihare. Considering the numbers, one may wonder whether the absence of MNC could be due to children's lack of ability to produce three term utterances. Chung's (1994) data from word order patterns show that around the age of 2, Korean children reach 30% of three term utterances, suggesting that lack of MNC is not related to inability to produce enough. Next possibility that we want to consider then is the possibility that the lack of MNC in the child speech may be related to some distributional patterns observed in the input. We take up this in the following section.

2.6.2. Input Patterns of MNC in Korean

The acquisitional finding on case markers showed that children remarkably match their input patterns. Similarly, we may consider the possibility that the lack of MNC in the child speech may be the influence of input patterns rather than the lack of some grammatical property involved in MNC. We examine Korean data, Jiwon's corpus from CHILDES (MacWhinny 2005) to check this possibility, assuming that the Japanese input will not be so drastically different from Korean input in the relevant respect.

We first searched the utterances with stative predicates which are compatible with the ones searched in Matsuoka's study. The results of the search are shown in table 9.

(30) Table 9. Input frequencies of psych-predicates with MNC/Nom objects

Stative predicates	Nom object	MNC
24	9	0

There were 24 utterances with such stative predicates. However, no instance of multiple Nom case was found. 9 out of 24 utterances were single Nom instances, in which Nom case appears only on the object. With respect to the types of the subject, some appeared with Topic marker, and some appeared as null subjects. But, surprisingly, there was no instance of Dative marked subjects. There was some instance of Dative case used with existential interpretation, but never to mark an experiencer. Given the greater frequency of Nom objects than that of MNC, it does not seem surprising that children in Matsuoka's study did not produce any MNC. Based on this fact, one may simply assume that children in her study may have reproduced their input pattern.

Next, we searched stative predicates that were not included in Matsuoka's study. Those are potential MNC instances with inalienable relation as in (8a). The results below show the input frequencies of all types of stative predicates and MNC, including those MNC with inalienable relation.

(31) Table 10. Input frequency of all types of stative predicates and MNC

Stative predicates	Single Nom	Single Nom
69	37	7

There were 69 utterances with the stative predicates. 7 instances were MNC utterances. Utterances with single Nom were 37, which was about 53% out of all stative predicates. Some examples are given below.

(32) a. *kkachi-ga ulgul-i ippe?* (MNC)

bird-Nom face-Nom pretty

'Does the bird have a pretty face?'

b. *Jini-nun aki-tta meri-ga uppessne* (Nom object w/ Topic Subject)

J-Top infancy-at hair-Nom not existed

‘Jin did not have hair at infancy.’

c. pro nun-i appa? (Nom object w/Null Subject)

eye-Nom hurt

‘Are you hurt in your eye?’

d. jiwoni-ga mul meko sipunde (Nom subject)

J-Nom water drink-NM want

‘Jiwon wants to drink water...’

Let us consider what this distribution of input tells us about the child grammar. It shows that the child seems to be getting some input of MNC to the extent that they receive about 10% of instances out of all utterances with stative predicates.¹⁵ Since Matsuoka did not include all types of stative predicates in her search, we do not know whether those children in her study ever produced any MNC of other types of stative predicates, besides MNC with psych-predicates. The child data in Matsuoka’s study and the Korean input in Jiwon’s corpus together show that it is not clear at this point whether the lack of MNC in the child speech in Matsoka’s study is really due to their lack of some grammatical knowledge or a reflection of input patterns. This situation indirectly suggests that Matsuoka’s finding alone may not necessarily argue for parameter-based account for the absence of MNC in Japanese children’s speech.

2.6.3. Learnability of MNC

We have considered two different accounts for the lack of MNC in the child speech: the input pattern or the lack of grammatical property of MNC. The production data we observed was not conclusive. Given that production data cover only what the child utters spontaneously, there is always a possibility that the data does not reveal what they could have produced. In other words, there is still the possibility that the child actually had the knowledge of MNC, but simply chose

¹⁵ 7.6% is the rate of MNC out of all utterances with stative predicates.

not to produce them because they prefer to express their intension using topic subject or null subject, instead of using Nom marked subject. Comprehension task can avoid this possibility. If the child has the relevant knowledge in their mental grammar, they should be able to provide appropriate interpretation for the sentences.

We design a comprehension task to tease apart different accounts for the lack of MNC in the child speech. We include DS sentences in the experiment to compare with MNC because DS can be distinguished from MNC by the lack of multiplicity of Nom as a sole difference. MNC and DS share the same predicates and Nom objects, but only differ with the case of the subject. We use MNC with inalienable relation because MNC of that type has a greater frequency than DS. If frequency indeed determines the order of grammar acquisition, we should expect that MNC will be acquired earlier than DS, given the greater frequency of MNC. However, if learning is guided by some grammatical factors, not just by sheer frequency, we expect some systematic discrepancy between the input patterns and what the child hypothesizes, which would require the mediation of UG.

We have two syntactic theories to consider for the possible discrepancy. First, according to Ura's multiple feature checking parameter, MNC is distinguished from DS by having positive setting of the multiple feature checking parameter. MNC is expected to be acquired later than DS because the child must set the positive setting of the parameter. According to Fukui and Takano's (1996) theory, the availability of the overt Nom case particle is a determining factor for the grammar which chooses Spell-out option over case checking. MNC follows as a natural consequence of choosing Spell-out option, and hence there is no difference between MNC and single Nom instance. This theory predicts that MNC will be acquired around the same time as DS.

In summary, the input-based approach predicts that MNC would be acquired earlier than DS given the rare instance of DS. The Default case approach predicts that MNC would be acquired around the same time as DS. Distinctively from these two approaches, the Multiple feature checking approach makes a prediction that MNC will be acquired significantly later than

DSC because the availability of MNC requires a positive setting of the multiple feature checking parameter. In order to test these different hypotheses, we conducted an experiment, which is the topic of the next section.

2.7. Experiment

2.7.1. Experimental Design

The experiment aimed to determine the relative order of acquisition of DS and MNC. A comprehension task is used to avoid the problem with production. We adopt Truth Value Judgment task (Crain and Mckee 1984) for a specific experimental method. An experimenter tells the child stories, and asks a puppet questions about the story. The child is asked to judge whether the puppet's answer is true or false based on the story by giving him Donut for the correct answer, or Peach for the wrong answer.

In constructing items to test the knowledge of the constructions, we consider one factor; that is the possibility that a given sentence can be understood by basic word order without the actual knowledge of case markers. In order to ensure that the child does not use the basic word order, we use the following strategy. Suppose that a sentence *S* is ambiguous between two readings, the target reading and the non-target reading, in a given context. If the child does not accept one reading consistently, even when the reading is very salient, we can conclude that she has not acquired the property responsible for that reading. Along this reasoning, each test sentence is made structurally ambiguous by taking advantage of head-final word order. In head-final languages, arguments of the matrix clause appear consecutively with arguments of the embedded clause, leaving the verb to be final. This allows the possibility that some argument potentially can be interpreted as a part of either the matrix clause or the embedded clause.

Let us see concrete examples. First consider a sample for MNC test. This test uses cleft constructions to create ambiguity between MNC reading and Non-MNC reading.¹⁶ This technique was first used in Sugisaki (2002).

(33) Sample story for MNC test

A cow, a dwarf, and a rabbit finds a small bed. They decide that the shortest one should take the bed. The rabbit says to the cow, “You have a fairly short height”. The dwarf says to the cow, “I think you have the shortest height among us”. But then, the rabbit says to the dwarf, “No, you have the shortest height.”

(34) MNC test item

nanjangi-ga kjang ki-ga jak-tako malhan-kes-un nwuku-n
dwarf-Nom most height-Nom short-comp said-comp-top who-Q

Under the story in (33), question (34) is structurally ambiguous between two readings depending on how to construe *nanjangi-ga* (dwarf). When it is considered as a part of embedded clause, the embedded clause forms MNC, as represented in (35a), rendering the whole sentence to be a question of the matrix subject. On the other hand, when it gets interpreted as a part of the matrix clause, as in (35b), the sentence yields a question for indirect object of the matrix verb.

(35) a. MNC reading

t_i [nanjangi-ga kjang ki-ga jak-tako] malhan-kes-un nwuku_i-ni?
dwarf-Nom most height-Nom small-comp said-comp-top who-Q
“Who_i was it that t_i said Dwarf has the shortest height?”

¹⁶ The reason that cleft construction was used to create the ambiguity is because Korean cleft construction allows the element in focus position to drop case markers, making more than one construal possible.

b. Non-MNC reading

nanjangi-ga t_i [kajang ki-ga jak-tako] malhan-kes-un nwuku_i-ni?

dwarf-Nom most height-Nom short-comp said-comp-top who-Q

“To whom_i was it that Dwarf said t_i, “Your height is the shortest?””

This is possible because *nwuku* in the focused phrase does not specify any case. Accordingly, there are two possible answers available: *rabbit* for MNC reading and *cow* for non-MNC reading. Next, consider a sample of DS test. The DS test item takes the simple form of two clause level wh-questions.

(36) Sample story for DS test

A duck and Youngja decide to run. The duck says to Youngja, “But, I don’t have running shoes”. Youngja says to the duck, “You don’t? But I have running shoes”. Then, mouse comes over, and says, “I have running shoes, too”. So Youngja and the mouse went running. Youngja comes home, and says to her mom, “Mom, the mouse has running shoes, but the duck doesn’t. Poor duck.”

(37) DS test item

Youngja-ga nwuku-ekey undongwha-ga iss-tako malhaess-ni?

Y-Nom who-to runningshoes-nom be-comp said-Q

Under the given story, question (37) is structurally ambiguous depending on where *nwuku-ekey* gets interpreted. If *nwuku-ekey* is interpreted as a part of embedded clause, the embedded clause forms DS, as shown in (38a). As a result, the whole sentence yields a question for the embedded subject. If *nwuku-ekey* is interpreted as a part of the matrix clause, the sentence yields a question for the indirect object of the matrix clause, as in (38b).

(38) a. DS reading

Youngja-ga [nwuku-ekey undongwha-ga iss-tako] malhaess-ni?

Y-Nom who-to runningshoes-nom be-comp said-Q

“Who_i did Youngja say [t_i has running shoes]?”

b. Non-DS reading

Youngja-ga nwuku-ekey [undongwha-ga iss-tako] malhaess-ni?

Y-Nom who-to runningshoes-Nom be-comp said-Q

“To whom_i did Youngja say t_i [that she has running shoes]?”

Two possible answers are available depending on each reading: *Mouse* for DS reading and *Duck* for non-DS reading.

16 Korean monolingual children with the age in range from 3;4 to 4;10 participated in the actual task. The experiment consists of 3 parts: training, pretest and the actual test. Training part started with a short instruction for the task disguised as a game. Three stories with each story followed by two simple questions were given for practice. Then, pretest was given in order to ensure that children have the knowledge of interpreting the similar sentence structures as the ones used in the actual task. There may be a possibility that the child may fail the test because she does not understand cleft constructions or one clause level wh-questions rather than the grammatical properties under investigation. There were four pretest stories, each of which was followed by 2 questions of the similar sentence structures as 2 types of test items. Only those who passed the pretest were included in the actual task. There were 6 test stories. In 4 out of the 6 stories, the puppet responds with target reading, and in the remaining 2 stories, the response was non-target reading. For all stories, a filler question was given along with the target question.

2.7.2. Predictions

This test permits three distinct predictions. If the acquisition of MNC and DS is determined by input frequency, MNC will be acquired before DS, predicting that there will be no children who pass DS, but fail MNC, as shown in (39). If MNC is a property of Nom being a generalized default case, MNC will be acquired around the same time. This approach predicts that children may pass both or fail both, rather than showing divergent behaviors on each test, as in (40a). On the other hand, if MNC is a property of multiple feature parameter setting, MNC will be acquired later than DS, since it takes time to reset the parameter from negative to positive value. So we do not expect to find the same children who fail DS would pass MNC, as shown in contingency table (40b).

(39) Input-driven learning

		<u>MNC</u>	
		Pass	Fail
<u>DS</u>	Pass	√	*
	Fail	√	√

(40) UG-based learning

a. Default case approach

		<u>MNC</u>	
		Pass	Fail
<u>DS</u>	Pass	√	*
	Fail	*	√

b. Multiple feature checking approach

		<u>MNC</u>	
		Pass	Fail
<u>DS</u>	Pass	√	√
	Fail	*	√

2.7.3. Results and Discussion

Children were classified as pass or fail, according to the pass/fail criterion. Children who gave 3 acceptances of target reading out of 4 chances were classified as passing. Otherwise children were classified as failing. The results are given below (See appendix for individual responses).¹⁷

(41) a. MNC vs DS test

		<u>MNC</u>	
		Pass	Fail
<u>DS</u>	Pass	11	5
	Fail	0	0

The results show that 6 children failed MNC, but all passed in DS, indicating that MNC is acquired later than DS. This experimental result cannot be explained by input-driven learning; if input-frequency was the determining factor, children should have acquired MNC earlier than DS given that they hear significantly more instances of MNC.

This result is consistent with Ura's (1996) theory of MNC. If MNC is an instance of a parameter setting, and if children have to undergo parameter resetting to allow multiple Nom case, the time required for the transition of grammar can account for the later acquisition of MNC. The nature of Nom case has been a long-standing issue in Korean and Japanese grammar studies. The syntactic analyses proposed for this issue have remained as theory internal questions without much empirical significance. Due to the lack of empirical evidence, the debate has been considered almost trivial even if they clearly reflect drastically different ways of looking at the nature of Nom case particle and consequently at their assumptions on what aspect of grammar allows parametric variation.

¹⁷ They are responses for target readings.

The present study took up an acquisitional perspective to shed light on this debate. We investigated the acquisitional implications of the two theories. The examination of the time course of acquisition for MNC vs DS shows that that multiple Nom case is not a property of default case, but rather an instance of a parameter setting of functional head T, indirectly suggesting that Nom case requires structural licensing.

2.8. Conclusion

In this chapter, we investigated the acquisition process of multiple case marking sentences in Korean. We observed from the spontaneous speech data that children do not produce MNC utterances even after they have learned single Nom case particle and the predicates with Nom objects. We considered whether this apparent lack of MNC is related to distributions in the input or something about the grammatical nature of MNC. In order to investigate this question, we considered grammatical theories of single Nom case and multiple Nom case marking property and analyzed their acquisition facts.

Concerning the acquisition order of Nom case, we observed that children produced Nom case earlier than any other case markers. We found from the study on input patterns that Nom case particle was the most frequently produced case in the input. We also observed that tense morpheme appeared around the same time as Nom case. Similarly, when we examined the input frequency of past tense morphemes, we found that they were also as frequent as Nom case, leading us to the possibility that the apparent close tie between the emergence of Nom case and the past tense morpheme may be related to input frequencies rather than the abstract grammatical licensing relation.

We turned to the developmental contour of Nom case in comparison with Acc case. We observed that children start with low frequency, and gradually converge on the adult-level of frequency for both Nom and Acc case. However, a difference was found in their developmental

pattern. Acc case was gradually increased while Nom case showed a relatively steep increase. Another surprising fact was that even if the input of Acc was very inconsistent due to the optionality, providing below 30% on average, the child acquires it as a fully grammatical option, and further successfully matches the input frequency in a precise manner. We discussed this phenomenon together with another case of input inconsistency observed in deaf children's acquisition of verbs of motion. Ross and Newport (1996) reported that deaf children who were exposed to inconsistent verbal morphemes (with 67% accuracy) somehow 'regularized the pattern', increasing the proportion up to 90%. This 'creolization' effect stands in a stark contrast with the Korean children's learning of Acc case. Why did the Korean children increase the frequency of Acc case? We were confronted with a question: how does the child know when to regularize and when not to?

This situation led us to suggest that children must have a prior knowledge to distinguish the two types of grammatical aspects. They should know what aspect of grammar can be optional or obligatory. If they know that Acc case can be optional, this prepares the child to stay with the input frequency. If they have the prior knowledge that verb agreement should be obligatory, they will increase the frequency beyond the level of input. The quick increase of Nom case can be explained along the same line of reasoning. The child must know that Nom case is obligatory, and that prior knowledge led them to produce Nom in all obligatory contexts, without going through gradual period.¹⁸

Concerning the acquisition of MNC, we started with Matsuoka's study on the acquisition of MNC with Japanese children. Matsuoka suggested that the lack of MNC in the child speech indicates that children acquire MNC later than single Nom case, supporting Ura's multiple feature checking parameter. We learned from examining input pattern in Jiwon's data that the input has very fewer instances of MNC than single Nom case, showing a similar pattern as in the child

¹⁸ The nature of the division between Nom and Acc nature with respect to the optionality remains to be seen. It could be related to the link between the abstract Nom Case and the EPP, as argued in Bošković (2002).

speech. In order to determine whether the lack of MNC was related to input pattern, or due to some grammatical factor, we conducted an experiment with a comprehension task. In the experiment we compare MNC with DS, considering the fact that the sole difference between DS and MNC is the multiplicity of Nom case.

Three different predictions were tested regarding the order of acquisition of MNC and DS. If the input frequency determines the order of the acquisition, MNC is expected to be acquired earlier than DS because there was a greater number of MNC than DS. Ura's theory predicts that MNC will be acquired later than DS; MNC requires learning of a positive value of a parameter. On the other hand, Fukui and Takano's theory predicts that MNC and DS will be acquired around the same time given that MNC stems from the default nature of a single Nom case. The results showed that children acquire DS earlier than MNC, supporting Ura's (1996, 2002) multiple feature checking parameter, in agreement with Matsuoka's (1998) original conclusion.

Appendix to Chapter 2

A. Individual Responses

	Age	D1	D2	D3	D4	N1	N2	N3	N4
1	4;0	Y	Y	Y	Y	Y	Y	Y	Y
2	3;11	Y	Y	Y	Y	Y	Y	Y	Y
3	3;7	Y	Y	Y	Y	Y	Y	Y	Y
4	4;1	Y	Y	Y	Y	Y	Y	Y	Y
5	4;10	Y	Y	Y	Y	Y	Y	Y	Y
6	4;0	Y	Y	Y	Y	Y	Y	Y	Y
7	4;8	Y	Y	Y	Y	Y	Y	Y	N
8	4;5	Y	Y	Y	Y	Y	Y	Y	Y
9	4;2	Y	Y	Y	Y	Y	Y	Y	Y
10	3;7	Y	Y	Y	Y	Y	Y	Y	Y
11	3;11	Y	Y	Y	Y	Y	Y	Y	Y
12	3;11	Y	Y	Y	Y	N	N	N	N
13	4;1	Y	Y	Y	Y	N	N	N	Y
14	3;8	Y	Y	Y	Y	N	N	N	Y
15	3;4	Y	Y	Y	Y	N	N	N	N
16	3;10	Y	Y	Y	Y	N	Y	N	N

D: Dative subject test

N: Multiple Nom case test

B. Test items

1) DS test

a. Story 1

A duck and Youngja decide to run. The duck says to Youngja, "But, I don't have running shoes". Youngja says to the duck, "You don't? But I have running shoes". Then, mouse comes over, and says, "I have running shoes, too". So Youngja and the mouse went running. Youngja comes home, and says to her mom, "Mom, the mouse has running shoes, but the duck doesn't. Poor duck."

Question

Youngja-ga nwuku-ekey undongwha-ga iss-tako malhaess-ni
Y-Nom who-to runningshoes-nom be-comp said-Q

b. Story 2

A dog and a tiger are playing. They both get tired and hungry after a while. The dog says to the tiger, 'I have nothing to eat.' Then, the tiger says to the dog, 'I have some grapes'. Then, a cat comes along, and says, 'I have some grapes, too.' Then, a big cow approaches, and asks, "Who has grapes?" The tiger quickly hides his grapes behind him, and answers, 'The cat has grapes, but I don't.'

Question

horangi-ga nwuku-ekey podo-ga iss-tako malhaess-ni?
tiger-Nom who-to grapes-Nom be-comp said-Q

c. Story 3

Sunhee finds a treasure box. When she opens it, she finds two coins inside. She takes one coin. Then, a cat comes along. Sunhee says to the cat, 'I have a coin.' The cat sees the treasure box, and finds one remaining coin. He takes the coin. A horse approaches. The cat says to the horse, 'I have a coin.' The horse sees the treasure box, and opens it. But he does not find any coin. Finally, the owner of the treasure box returns and finds that his two coins are gone. He asks, "Where are my coins?" The cat quickly hides his coin, and says, 'Sunhee has a coin, but I don't.'

Question

koyangi-ga nwuku-ekey dongcen-i isstako malhaess-ni?
cat-Nom who-to coin-Nom be-comp said-Q

d. Story 4

It is Christmas season. Santa prepares two pieces of wood, and goes somewhere. A farmer comes along, and takes one piece. A mouse approaches. The farmer says to the mouse, 'I have a piece of wood.' The mouse sees the other wood, and takes it. Then, a deer approaches. The mouse says to the deer, 'I have a piece of wood.' The deer looks around, but he does not find any. Finally, Santa returns, and finds that all his wood is gone.' The mouse quickly hides his wood, and says, 'The farmer has wood, but I don't.'

Question

sangchi-ga nwuku-ekey namu-ga iss-tako malhaess-ni?
mouse-Nom who-to wood-Nom be-comp said-Q

2) MNC test

a. Story 1

A cow, a dwarf, and a rabbit find a small bed. They decide that the shortest one should take the bed. The rabbit says to the cow, "You have a fairly short height". The dwarf says to the cow, "I think you have the shortest height among us". But then, the rabbit says to the dwarf, "No, you have the shortest height."

Question

nanjangi-ga kajang ki-ga jak-tako malhan-kes-un nwuku-ni
dwarf-Nom most height-Nom short-comp said-comp-top who-Q

b. Story 2

A pig is sleeping. A tiger, a dog and a sheep come along. They decide that the one who has the loudest voice should wake up the pig. The dog says to the sheep, "You have a fairly loud voice". Then, the tiger says to the sheep, "I think you have the loudest voice among us". But, then, the dog says to the tiger, "No, you have the loudest voice."

Question

horangi-ga kajang moksori-ga ku-tako malhan-kes-un nwuku-ni
tiger-Nom most voice-Nom loud-comp said-comp-top who-Q

c. Story 3

Jasmin, a little girl, and Wonderwoman are walking in the street. They find a beautiful comb. They decide that the one who has the longest hair should get the comb. A little girl says to Wonderwoman, "You have a fairly long hair. Then, Jasmin says to the Wonderwoman, "I think

you have the longest hair among us”. But then, the little girl says to Jasmin, “No, you have the longest hair.”

Question

Jasmin-i kajang meri-ga kil-tako malhan-kes-un nwuku-ni
J-Nom most hair-Nom long-comp said-comp-top who-Q

d. Story 4

An elephant, a horse, and a dog are walking in the wood. They find a little child crying. The child says, “I am lost.” They decide that the one who has the biggest shoulder should carry the child. The horse says to the dog, “You have a fairly big shoulder. Then, the elephant says to the dog, ‘I think you have the biggest shoulder among us’. But then, the horse says to elephant, “No, you have the biggest shoulder.”

Question

kokkiri-ga kajang akke-ga kkun-tako malhan-kes-un nwuku-ni
elephant-Nom most shoulder-Nom big-comp said-comp-top who-Q

Chapter 3. Scrambling Theories from Acquisitional Perspective

3.0. Introduction

This chapter and the following chapter investigate the acquisition of a grammatical phenomenon called ‘scrambling’, as illustrated in (1).

- (1) John-ul_i Mary-ga t_i milesstta.
J-Acc M-Nom pushed
‘Mary pushed John’

- (2) Mary-ga John-ul milesstta
M-Nom J-Acc pushed

Examples (1) and (2) differ from each other in the position of the object *John-ul*. Example (2) shows the canonical word order of Korean, SOV, with the object following the subject. In (1), the object *John-ul* appears in the sentence initial position, exemplifying scrambling of the object. Strictly speaking, this kind of object fronted surface word order alone does not tell us whether it is an instance of scrambling or not, since even a language like English allows the object to appear in the sentence initial position via topicalization, as shown in (3).

- (3) Beans, John likes.

However, in Korean, it is clear that (1) is an instance of a grammatical operation different from topicalization. Korean has a morpheme *-un*, which indicates topic or focus. When an object is

topicalized or focalized, this morpheme is attached to the object, replacing the Accusative case with the morpheme *-un*, as shown below.

(4) John-un Mary-ga milesstta.

J-Top/Focus M-Nom pushed

‘As for John, Mary pushed him.’

‘It is John that Mary pushed.’

This means that the word order change in (1) is not a result of topicalization or focalization, but something different. The operation in question is standardly referred to as scrambling.

It is known that languages vary as to whether this operation of scrambling is allowed or not. The cross-linguistic variation of scrambling raises an acquisitional question: how does the child acquire this language particular property? One may easily assume that the child learning a scrambling language would receive a robust amount of positive evidence from the input. However, to our surprise, we find that the input frequency of scrambling is almost close to zero, presenting a poverty of stimulus situation. This raises a learnability question. How does the child acquire scrambling under such impoverished input? To be more specific, what kind of evidence does the child rely on to acquire scrambling, then?

In this chapter we review major grammatical properties of scrambling, and discuss syntactic theories of scrambling from an acquisitional perspective. Three different analyses of scrambling are reviewed: the EPP/Focus theory by Miyagawa (2003), the Scrambling feature theory by Grewendorf and Sabel (1999) and the Case-related theory by Bošković (2004, in press). It will be shown that each theory makes different implications for several acquisitional issues. In section 3.1, we clarify the term ‘scrambling’ by discussing the defining characteristics of scrambling. Section 3.2 looks into two major syntactic properties of scrambling: optionality and its dual nature. Section 3.3 points out a correspondence between the landing site of the

scrambling and types of scrambling. Section 3.4 reviews three syntactic theories of scrambling, and discusses their implications for cross-linguistic variation. Section 4.5 discusses implications of each theory of scrambling for several acquisitional questions.

3.1. Scrambling

Descriptive grammar studies have classified languages into two groups according to how grammatical relations are expressed: word order languages and inflectional languages. Languages such as English are classified as word order languages, in which word order expresses which element is the subject, and which element is the object, etc. Languages such as Russian are categorized as inflectional languages. Free word order is often recognized as a property of inflectional languages. It is assumed that case inflections identify the grammatical roles of displaced elements, allowing words to be freely ordered.

The phenomenon of free word order has been termed as ‘scrambling’ since Ross’ (1967) seminal work. Ross analyzed scrambling as a movement operation. For instance, for the following scrambling sentence, he proposed that the object *John-ul* is preposed to the sentence initial position via movement operation.

(5=1) John-ul_i Mary-ga t_i milessta.
 J-Acc M-Nom pushed

In the introduction, we pointed out that scrambling can be distinguished from topicalization or focalization by the morphological marker attached to the scrambled element. The topic or focus element carries the overt marker *-un*, while the scrambled element keeps the Acc case marker. This is not the case for many other scrambling languages. Many scrambling languages do not have overt topic or focus marker. For instance, Russian and Serbo-Croatian are scrambling languages, but lack an overt topic or focus marker. This means that in these languages a single

object fronting to the beginning of a sentence would be ambiguous between scrambling and topicalization or focalization.

In these languages, further grammatical characteristics allow us to distinguish scrambling from other types of displacement operations.¹ One of them is the property of ‘multiple scrambling’; scrambling allows displacement of multiple elements, while topicalization or focalization cannot operate on more than one element, as shown below.

(6) a. A book_i John gave Mary t_i. Topicalization

b. *A book_j Mary_i John gave t_i t_j.

(7) a. A book_j Mary-ekey_i John-i t_i t_j cwuesstta. Scrambling

book-Acc M-dat J-Nom gave.

‘John gave Mary a book.’

b. *chaek-un_j Mary-nun_i John-i t_i t_j cwuesstta. Topicalization/Focalization

book-Top/Focus Mary-Top/Focus John-i gave

(8) Tamil (Sarma 1999)²

a. Mary-kku_i oru book-ai_j John t_i t_j kodothan Scrambling

M-Dat one book-Acc John gave

‘John gave a book to Mary.’

b. *Mary-kku-naa_i oru book-ai-naa_j John t_i t_j kodothan Topicalization

Mary-Dat-Top one book-Acc-Top John gave

¹ See section 3.2 for detailed discussion of syntactic properties of scrambling.

² Unlike Korean/Japanese, Tamil topic marker ‘-naa’ does not replace the case marker. Instead, it is attached after the case marker.

The single object fronting in (6a) is grammatical, being an instance of topicalization. The ungrammaticality of (6b) indicates that topicalization does not allow dislocations of multiple elements, and further that English does not have scrambling, which would have been produced such a derivation. We observe that topicalization in Korean and Tamil exhibits the same constraint as English. (7b) and (8b) both involve multiple topic/focus; consequently, they are judged ungrammatical. On the contrary, (7a) and (8a), where both the direct object and the indirect object are scrambled, are judged grammatical, indicating that the property of multiple displacements is peculiar to scrambling.

Another distinct syntactic property of scrambling is its semantic vacuity. Based on the following example from Japanese, Saito (1989) observes that scrambling does not have semantic import in the sense that it does not create an operator-variable relation.³

(9) a. Nani-o_i John-ga [Mary-ga t_i katta ka] sitteiru.

What-Acc J-Nom M-Nom bought Q knows

‘John knows what Mary bought.’

b. [Mary-ga nani-o katta to]_i John-ga Bill-ga t_i itta ka sitteiru.

M-Nom what-Acc bought that John-Nom Bill-Nom said Q knows

‘John knows what Bill said that Mary bought.’

Example (9a) is a declarative sentence containing an interrogative embedded clause. The embedded object wh-phrase is scrambled to the beginning of the sentence. One may expect that this should cause ungrammaticality, since the wh-phrase is located outside of the clause in which it takes scope. However, the sentence is still judged grammatical, with the wh-phrase taking embedded scope. This means that scrambling does not establish an operator-variable relation, and

³ The relevant data involves long distance scrambling, in which scrambling takes place out of the embedded clause to the matrix clause.

hence can be undone in LF so that the *wh*-phrase is within its scope in LF. (9b) involves scrambling of the most embedded CP containing a *wh*-phrase to the matrix clause. The *wh*-phrase is again outside of its scope, the intermediate CP. However, the sentence is still judged grammatical, again indicating that the scrambled phrase does not create an operator-variable relation, and hence can be put back in LF. The same fact is observed in Korean and other scrambling languages (See Bošković 2004). Bošković and Takahashi (1998) demonstrate that topicalization differs from scrambling in this respect. Consider the example below.

(10) *[That picture of *who*]_i, I know who bought *t_i*

In (10), a phrase containing the *wh*-phrase is topicalized, similar to the case of scrambling in (9b).⁴ The ungrammatical status shows that *who* cannot take embedded scope once the phrase containing it is topicalized out of the embedded clause, unlike the case of scrambling. The topicalized phrase cannot be put back in LF, so that the *wh*-phrase remains outside of its scope, resulting in ungrammaticality. Based on this fact, Bošković and Takahashi argue that scrambling does not create an operator-variable relation, while topicalization does.

So far we have reviewed two defining characteristics of scrambling. Extensive studies of scrambling have revealed that the operation of scrambling exhibits very complex behavior. In the following section, we review some of the major findings.

3.2. Syntactic Properties of Scrambling

Since Ross' (1967) seminal work, extensive research has been done concerning the phenomenon of scrambling. Ross (1967) first took scrambling as a kind of movement operation, and analyzed it as a stylistic movement, implying that it does not exhibit any syntactic properties.

⁴ We cannot run the test in (9a) with respect to topicalization because *wh*-phrases in English cannot be topicalized.

Subsequent research across languages, however, has identified various syntactic properties that are associated with scrambling (Mahajan 1990 for Hindi, Karimi (2003) for Persian, Choi (1994) for Korean, Saito (1985) for Japanese, Bailyn (2001) for Russian, Stjepanović (1996) for Serbo-Croatian, to name a few).

3.2.1. Optionality and Dual properties

Two major properties of scrambling that have received considerable focus are optionality and dual properties. Scrambling is optional in that the operation is not required in any sense. Canonical order is always available as an alternative. In Government Binding framework (Chomsky 1981, Lasnik and Saito 1988), optionality was not a problem. Move α allows movement anywhere and anytime, as long as it does not violate other constraints formulated as principles such as 'Case Filter' or 'Projection Principle'. However, the optionality has become a challenge to the syntactic theory of Minimalist Program (Chomsky 1994). In the Minimalist theory syntactic movement takes place only for a reason in the consideration of economy, particularly, for the purpose of morphological feature checking. This means that when there is no need of checking any morphological feature, no syntactic operation is allowed. For instance, the raising of *that man* in (11b) is not allowed because the case feature of the NP is checked in- situ by the preposition *to*, while *that man* in (11a) is required to move to check its case feature because it cannot be case-checked in- situ.

(11) a. That man_i is t_i in the garden.

b. *That man_i seems to t_i that it is raining outside.

In the case of scrambling in Korean, given that unscrambled order SOV is always possible, without a semantic change from scrambled order (in the case of long-distance scrambling), scrambling seems unmotivated. From the point of movement being conceived as a

morphologically driven operation, the optionality of scrambling then becomes a violation of a core assumption of the Minimalist theory. Bošković and Takahashi (1998) argue for a base-generation approach to scrambling for this reason. They propose that scrambling is an instance of lexical insertion in a non-theta position. Other analyses of scrambling derive the optionality by postulating a morphological feature such as EPP feature (Miyagawa 2003) or a scrambling feature Σ (Grewendorf and Sabel 1999) as a driving force for movement. We will look into the details of the three approaches in the following section.

The second characteristic of scrambling is its dual nature. Research has established that phrasal movement operations can be classified into two types: A-movement and A-bar movement. For example, *wh*-movement is a typical type of A-bar movement, and object-raising in passive is a typical case of A-movement. Interestingly, scrambled elements have been shown to exhibit both A-properties and A-bar properties. The A vs A-bar distinction can be tested by several syntactic diagnostics. A-bar movement properties include Weak Crossover effects, Reconstruction effects, and unboundedness. A-movement properties include the lack of above A-bar movement properties. Neutralization of Weak Crossover effects, the ability to bind an anaphor and clause-boundedness are the properties of A-movement. In what follows, we show how scrambling behaves with respect to each of these properties.

First, let us consider Weak Crossover effect. Sentences in (12) illustrate the difference between A-movement and A-bar movement with respect to the Weak Crossover effect (Corver and Riemsdijk 1994).

- (12) a. *Who_i does his_i mother dislike t_i?
 b. *Who_i does it seem to his_i mother that Mary dislikes t_i?
 c. John_i seems to his_i mother t_i to be smart.⁵

⁵ Notice that *wh*-question counterpart of (12c) given in (i) does not show Weak Crossover effect.
 (i) Who_i seems to his_i mother t_i to be smart?

In (12a/b), the *wh*-operator *who* binds bound variable pronoun *his* and the trace at the same time, yielding a Weak Crossover configuration.⁶ The same holds for *John* in (12c); It binds both the pronoun *his* and the trace. The coreference of *who* with the pronoun is not allowed in (12a/b), while the coreference of *John* with the pronoun is allowed in (12c). This contrast has been attributed to the type of movement involved in each case. Cases of (12a/b), involving movement to an A-bar position (A-bar movement), result in a Weak Crossover effect, while (12c), involving movement to an A-position (A-movement), is free from such effect. Now, let us look at the case of scrambling. Consider the contrast between (13a), in which the object is in-situ, and (13b), in which the object is scrambled to the sentence initial position.

(13) a. * *ku-uy_i apeci-ka nwuku-lul_i sileha-ni?*

he-Gen father-Nom who-Acc dislike-Q

‘Who_i does his_i father dislike?’

b. *nwuku-ul_i ku-uy_i apeci-ka t_i sileha-ni?*

who-Acc he-Gen father-Nom dislike-Q

If we assume that the in-situ *who* in (13) undergoes covert operator movement, and binds the pronoun and the trace at the same time, we can account for the ungrammaticality of the sentence as a Weak Crossover effect, on a par with (12a/b) above. To the extent that this is true, the grammaticality of (13b), which involves scrambling of the *wh*-operator *nwuku-ul* to the pre-subject position, suggests that scrambling exhibits A-properties with respect to the Weak Crossover effect (Choi 1994).

This is not a counterexample for lack of Weak Crossover effect for A-bar movement. This fact is considered as indicating that it is the A-movement prior to A-bar movement that neutralizes Weak Crossover effect.

⁶ This way of defining the configuration for the Weak Crossover effect is proposed by Koopman and Sportiche (1982) in their ‘Bijection Principle’, as stated below.

(i) Every variable must be bound by exactly one operator.

(ii) Every operator must bind exactly one variable.

Besides the lack of the Weak Crossover effect, the ability to bind an anaphor indicates another A-movement property of scrambling. The contrast in (14) shows that A-movement opens up a new binding possibility, whereas topicalization, an instance of A-bar movement, does not.

- (14) a. John and Mary_i seem to each other t_i to be smart.
 b. *John and Mary_i, each other do not really like t_i.

Consider the case of scrambling below. (15a) is excluded as a violation of Binding Condition A because the reciprocal is not bound. However, the scrambling counterpart, in which the object is preposed to the clause initial position, as in (15b), is grammatical, indicating that scrambling creates a new binding possibility for anaphors, as in the A-movement case (Choi 1994).

- (15) a. *selo- uy_i chinku-ka kutul-ul_i kosohaysstta.
 Each other -Gen friend-Nom they-Acc sued
 'Each other's friends sued them.'
 b. kutul-ul_i selo-uy_i chinku-ka kosohaysstta.
 They-Acc each other-Gen friend-Nom sued.
 'Each other's friend sued them.'

Though scrambling seems to be an instance of A-movement from the phenomena discussed so far, with respect to other diagnostics such as Reconstruction effect and unboundedness, it behaves like A-bar movement. First, consider the following contrast between A-movement and A-bar movement with respect to a reconstruction possibility (Choi 1994).

- (16) a. Himself_i, John like t_i.
 b. *Himself_i seem to John t_i to be intelligent.

The grammaticality of (16a) shows that the anaphor *himself*, which has undergone topicalization, a type of A-bar movement, can be reconstructed to be bound by *John*. The opposite situation is observed in (16b). *Himself*, which has undergone A-movement, cannot be bound, resulting in a violation of Binding Condition A. The contrast indicates that an element which moves to A-bar position, but not A-position, can be reconstructed. Let us see how scrambling behaves with respect to reconstruction effects. Consider (17), which employs scrambling of an anaphor.

(17) kucasin-ul_i John-i t_i binanhasstta.

‘himself-Acc J-Nom criticized

‘He_i criticized himself_i.’

Its grammatical status demonstrates that the scrambled anaphor can be reconstructed, as if it is an A-bar movement (Choi 1994).

Another A-bar movement property of scrambling is its unboundedness. It is standardly assumed that A-movement is clause-bounded (i.e., It cannot take place out of finite clause). This is shown in the contrast between (18a) and (18b).

(18) a. *John_i seems t_i likes Mary.

b. John_i seems [t_i to like Mary].

(19) a. Who_i does it seem t_i likes Mary?

b. Who_i does John seem [to like t_i] ?

(20) nwuk-ul_i John-i Mary-ga t_i pulses-tako malhaess-ni?

who-Acc J-Nom M-Nom called-Comp said-Q

‘Who_i did John say that Mary called t_i?’

A-movement of *John* is fine in (18b) because the embedded clause from which *John* has moved out is non-finite. The same movement of *John* is blocked in (18a) because the embedded clause is finite. This finiteness distinction in movement does not appear in the case of A-bar movement, as shown in (19). A-bar movement out of both finite clause, as in (19a) and non-finite clause, as in (19b), is possible. Scrambling behaves as if it were A-bar movement, allowing long-distance application, as shown in (20).

An additional peculiar property of scrambling is the argument vs adjunct asymmetry, known as the ban on long-distance scrambling of adjuncts. The behavior of scrambling with respect to the argument vs adjunct distinction is illustrated in (21) below.

- (21) a. nwuk-ul_i John-i Mary-ga t_i pulses-tako malhaess-ni?
 who-Acc J-Nom M-Nom called-Comp said-Q
 ‘Who did John say that Mary called t?’
- b. ?? ppali_i John-i [Mary-ga t_i pab-ul mekess-ko] malhasstta.
 quicky J-Nom M-Nom mean-Acc ate-Comp said
 ‘John said that Mary ate the meal quickly.’

The degraded grammaticality status of (21b) shows that the adjunct can not be interpreted as modifying the embedded predicate, indicating that the adjunct cannot be scrambled, while there is no such restriction for argument scrambling, as in (21a). We can relate this asymmetry to scrambling as a type of A-movement, assuming that A-movement affects only arguments, but not adjuncts (Bošković and Takahashi 1998).

As we saw in the introduction, iterability (multiple scrambling) and semantic vacuity (in the case of long-distance scrambling) are further properties of scrambling. Both properties are less clear with respect to the issue of whether they indicate A-movement or A-bar movement. The properties of scrambling that have been discussed so far are listed below:

(22) Syntactic properties of scrambling

a. A-properties:

- Lack of Weak Crossover effect
- Anaphor binding
- Argument vs Adjunct Asymmetry

b. A-bar properties

- Reconstruction effect
- Unboundedness⁷

c. Neutral with respect to A vs A-bar properties

- Multiple scrambling
- Semantic vacuity

It has been noted that there is variation across scrambling languages as to which properties of the above are exhibited in a given scrambling language. Some languages exhibit both A-properties and A-bar properties, but some languages only A-bar properties. Scrambling languages such as Korean, Japanese (Saito 1985), Tamil (Sarma 1999), Hindi (Marhajan 1990), Russian (Bailyn 2000), German (Haeberli 2003), and Persian (Karimi 2003) have been reported to show both A-scrambling and A-bar scrambling properties.⁸ On the other hand, Turkish (Kornfilt 1994) and Serbo-Croatian (Bošković and Takahashi 1999) are known to show only A-bar properties.

⁷ German has been traditionally considered as a scrambling language, but does not allow long-distance scrambling, as in (i) below.

(i) *daß keiner Hygrometer sagt daß Antje mag
that no-one hygrometers says that Antje likes
'no one says that Antje likes hygrometers.'

Further, unlike other scrambling languages, German does not allow scrambling of wh-words (Müller 1995) and its scrambling is claimed to have always semantic effects. This divergence has led to different views on the status of German with respect to scrambling. Grewendorf and Sabel (1999) classify German as an A-bar scrambling language, distinguishing it from Japanese type of scrambling, which they consider as an A-scrambling language. Bošković (2003) does not consider German as a scrambling language.

⁸ German and Russian are controversial on this issue. Grewendorf and Sabel (1999) argue that German shows only A-bar properties, but Haeberli (2002) shows that German does allow anaphor binding of the scrambled element.

To summarize, we have seen that scrambling is an optional operation, and has dual properties. There is cross-linguistic variation among scrambling languages concerning whether a given scrambling language exhibits both A-properties and A-bar properties, or only A-bar properties. Taking this into consideration, the following cross-linguistic typology of scrambling emerges.

(23) Table 1. Typology of Scrambling

Yes	No
<i>A-scrambling and A-bar scrambling</i> ----- Korean, Japanese, Hindi, Tamil, Persian, ... <hr/> <i>A-scrambling</i> <i>A-bar scrambling</i> * Turkish, Serbo-Croatian	English, Spanish, Navajo, Chinese, French, Thai, Vietnamese, Icelandic,

Languages are separated into two groups with respect to the phenomenon of scrambling: languages which allow scrambling and languages which do not. There is also some variation within scrambling languages as to which type of scrambling is exhibited. This variation among scrambling languages makes it difficult to define scrambling. In the present study we take the possibility of multiple scrambling and semantic vacuity as the defining properties of scrambling.

3.3. Pre-subject Scrambling: both A and A-bar properties

There is a further aspect of scrambling that adds to the complex behavior of the phenomenon. It concerns the interaction between the landing site of scrambling and the two types of scrambling. Mahajan (1990) made an important observation that scrambling shows different properties depending on the landing site of scrambling. He noted with respect to Hindi that scrambling to pre-subject position in a simple clause can be either A-movement or A-bar movement, but scrambling crossing a clause (long-distance scrambling) can only be an A-bar movement. The same observation has been made in other scrambling languages such as in

Japanese. For instance, Saito (1992:76) observed that the pre-subject scrambling in a simple clause shows A-property with respect to anaphor binding, but long-distance scrambling can exhibit only A-bar properties.

(24) a. *otagai-no_i sensei-ga karera-o_i hihansita koto
 each other-gen teacher –Nom they-Acc criticized fact
 ‘Each other_i’s teacher criticized them_i.’

b. karera-o_i otagai-no_i sensei-ga t_i hihansita koto
 they-Acc each other-gen teacher-Nom criticized fact

(25) a. *otagai-no_i sensei-ga [Hanako-ga karera-o_i hihansit-to] itta koto
 each other-Gen teacher-Nom H-Nom they-Acc criticized-Comp said fact
 ‘Each other_i’s teachers said that Hanako criticized them_i.’

b. *karera-o_i otagai-no_i sensei-ga [Hanako-ga t hihansita-to] itta koto
 they-Acc each other-Gen teacher-Nom H-Nom criticized-Comp said fact

Example (24a) is ill-formed because the reciprocal *otagai-no* is unbound, a violation of Binding Condition A. However, the reciprocal, which could not be bound in (24a), is bound by the scrambled pronoun in (24b), indicating that the short scrambling can create a new binding configuration. However, this ability is not replicated in the case of long-distance scrambling. The preposed scrambled pronoun to the clause-initial position across the embedded clause cannot bind the reciprocal in the subject position, as shown by the ill-formedness of (25a). This indicates that unlike the case of short distance scrambling, long-distance scrambling cannot not create a new binding configuration.

The grammatical properties of scrambling we observed so far raise some important acquisitional questions. First, how does the child come to know whether her language allows

scrambling or not? Secondly, how does the child find out which type of scrambling is employed in her language? Particularly, what kind of linguistic experience can serve as positive evidence for the availability of scrambling as well as each type of scrambling employed in her language? Before we explore these questions, we review syntactic theories that have been proposed to account for these scrambling properties.

3.4. Syntactic Theories of Scrambling

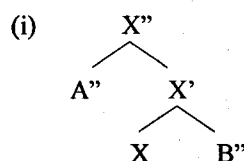
The early syntactic account of scrambling does not seem to be so different from the traditional view of free word order. Instead of classifying languages as word order languages vs inflectional languages, a notion of configurationality was adopted to capture the variation. It was hypothesized that languages differ in terms of structural configuration; some languages are structured configurationally, conforming to hierarchical structure, while other languages are not, allowing flat structure. Free word order was recognized as one of the major characteristic features of non-configurational languages (Hale 1982, Farmer 1980, Haider 1988).⁹ Hale (1980) attempted to derive the difference between configurational and non-configurational languages by parameterizing the notion of government in X-bar theory.¹⁰ He proposed that in non-scrambling

⁹ Hale (1982) observes that languages which allow scrambling often share the following characteristics:

- i. The use of discontinuous expressions.
 - ii. Free use of "pronoun drop"
 - iii. Lack of the NP-movement transformation
 - iv. Lack of pleonastic NPs (like it, there, il....)
 - v. Use of rich case system.
 - vi. Complex verb words or verb-com-Aux system.
- To this list, Huang (1982) adds one more characteristic:

- viii. The lack of standard ECP effects

¹⁰ Hale (1982:89) defines government as a relation which holds between the head of a category and its immediate sisters. In a configuration like (i) below, the category A'' is governed by X', and the category B'' is governed by X.

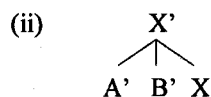


languages, government “clicks”, imposing a particular order of arguments (e.g. the object to the right, the subject to the left), while in scrambling languages, it “shuts down”, allowing an argument to appear on either side of the head in X-bar theory. Accordingly, scrambling was viewed as a result of base-generation of arguments on either side of the head.

This non-configurational analysis has been challenged by detailed studies of the structural properties in various scrambling languages. Fukui (1986) and Saito and Hoji (1983) provide convincing arguments for the existence of hierarchical structure in Japanese, with evidence from various tests which make reference to the hierarchical relation of elements. With the non-configurational view abandoned, an alternative approach was advanced to account for free word order phenomenon. It was argued that scrambling is a result of syntactic movement which is applied optionally (Hoji 1985, Saito 1985).

Scrambling as an optional movement was not a problem in a system with Move α , where movement can take place as long as it does not violate any principle. However, in the Minimalist theory (Chomsky 1995), according to which syntactic operations only take place for a reason, optionality has been recognized as a serious problem for the theory. Two different lines of research have been taken to derive the optionality. One line of approach is to postulate a morphological feature as the driving force for scrambling. Such an approach is taken by Miyagawa (2003) and Grewendorf and Sabel (1999). Miyagawa suggests that scrambling is an EPP or focus driven movement. Grewendorf and Sabel postulate scrambling feature Σ . The other line of approach is to assume that scrambling is actually not a movement, but the surface effect of merging arguments in non-theta positions. The latter approach is proposed by Bošković and

The above configurational structure allows government to distinguish nominal arguments of a lexical head; A” as a sister of X’ and B” as a sister of X. Non-configurational languages have a flat phrasal structure, as depicted below.



In the flat structure, government does not distinguish among nominal arguments because both are governed by X, or alternatively, government simply does not operate, as Hale suggests.

Takahashi (1998) and Bošković (2004, in press). In some sense, the latter view is a revival of Hale's original analysis of scrambling. However, it has several crucial differences, one important difference being that Bošković and Takahashi's base-generation approach analysis does not assume flat structure.

The three scrambling analyses also differ in the way they account for the dual property of scrambling. Miyagawa (2003) takes scrambling as a non-unitary phenomenon, and locates distinct sources for each type of scrambling: EPP feature for A-scrambling and Focus feature for A-bar scrambling. On the other hand, the two proposals, Grewendorf and Sabel's (1999) scrambling feature theory and Bošković's (2004, in press) case-related, take scrambling as a unitary phenomenon, and derive the dual nature by resorting to a language specific factor, the availability of multiple specifiers. The following sub-sections review these three syntactic analyses of scrambling in detail and their implications for cross-linguistic variation.

3.4.1. EPP/Focus Approach: Miyagawa (2003)

Miyagawa (2001, 2003) assumes that scrambling is a non-unitary phenomenon. He hypothesizes that the two types of scrambling are driven by two different morphological features: EPP feature for A-scrambling and Focus feature for A-bar scrambling. He argues that A-scrambling is conditioned by two factors: V-to-T raising and overt morphological case, as stated in (26).

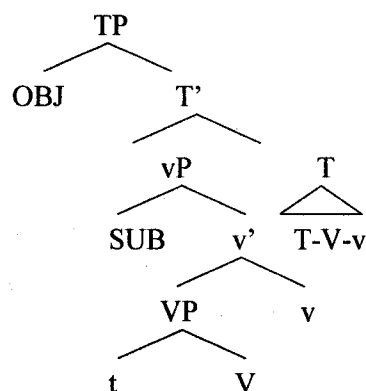
(26) Conditions for A-scrambling (Miyagawa 2003)

A language allows EPP-driven scrambling of the object if the language has V-to-T raising and overt morphological case marking.

V-to-T raising plays the role of extending the checking domain of T, making the object equally local to T as the subject. The presence of overt Acc morphological case allows the scrambled element to be licensed by the head T, by having an agreement relation with T.

The latter assumption is based on a particular view of overt morphological case that Miyagawa adopted. Following Kuroda's (1988) approach to morphological case marking, Miyagawa assumes that morphological cases differ from abstract cases in that any morphological case can undergo agreement relation with T, simply by virtue of being a morphological case. The agreement relation enables the object with Acc case marker to be licensed by the head T, allowing it to move to spec TP. The object is allowed to move to spec TP, as long as it has overt morphological case. The abstract Acc case, however, does not have this ability. It can only agree with the relevant position in a certain configuration, requiring the object to be licensed solely by the small v. When V-to-T raising and the morphological case are available, a given language allows the object to move to spec TP over the subject, resulting in the scrambling word order OSV. This is the nature of the A-scrambling operation. The schematic derivation goes as follows:

(27) A-scrambling of object



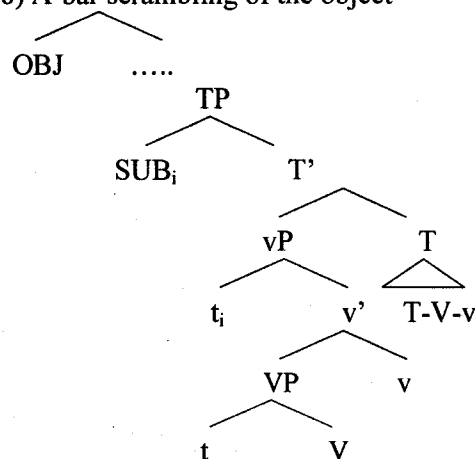
T has an EPP feature which needs to be checked. V raises to T, picking up v. At this point, Spec TP and Spec vP become equidistant from the object by virtue of V-to-T raising. The object, having overt morphological case marking, undergoes an agreement relation with the head T. This

allows the EPP feature on T to attract the object across the subject without violating locality, giving rise to the OSV surface word order.

Under this system, the same movement, EPP triggered operation, gives rise to both SOV and OSV orders. SOV is the instance in which the subject is moved to spec T for the checking of EPP-feature in T, instead of the object. A-scrambling is just a conspiracy of two independent language specific properties: V-to-T raising and overt morphological case marking. This system correctly predicts that French, a language with V-to-T raising uncontroversially, does not allow scrambling because it does not have overt morphological case. Even if V-to-I raising were to extend the checking domain, the object is not allowed to raise to spec TP since the object with abstract Acc case cannot agree with T. In a similar way, languages with overt case morphology may not allow A-scrambling if V-to-T raising is not available. In this case, the object will not be in the same checking domain as the subject.

Having discussed the mechanism of A-scrambling, let us see how this system derives A-bar scrambling. Miyagawa (2003) assumes that A-bar scrambling is a focus driven movement, in parallel with WH-movement. The scrambled phrase moves to a position above TP. The subject moves to spec T to check EPP feature of T, and the object moves to a higher position, a focus-related position, as schematized below.

(28) A-bar scrambling of the object



The EPP/Focus theory of scrambling seems to face two problems concerning cross-linguistic facts. First, since the theory argues that V-to-T raising and overt case morphology are the two conditions for scrambling, it predicts that Icelandic, which meets both conditions, will allow scrambling. However, that is not true. Scrambling is not possible in Icelandic. Another problem has to do with cross-linguistic variation with respect to A-scrambling vs A-bar scrambling. The theory assumes that there is no single factor which distinguishes scrambling languages from non-scrambling languages. Given that the triggering features for each type of scrambling are different, it is expected that languages may have one of three possible options: only A-scrambling or only A-bar scrambling, or a mixture of both types of scrambling. The limited cross-linguistic data available to us show that the majority of scrambling language seems to exhibit mixed properties.¹¹ Korean (Choi 1994), Japanese (Saito 1985), Persian (Karimi 2003), Hindi (Marhajan 1990), and Tamil (Sarma 1999) are reported to show both A and A-bar scrambling properties. Languages such as Serbo-Croatian (Bošković and Takahashi 1998), and Turkish (Kornfilt 1994) are known to have only A-bar scrambling properties. Interestingly, no scrambling language has been found to exhibit A-scrambling properties exclusively.¹² In this system, the lack of A-scrambling language remains as an accidental gap in the typological possibilities.

3.4.2. Scrambling Feature Approach: Grewendorf and Sabel (1999)

Grewendorf and Sabel (1999) consider scrambling as a unitary phenomenon, postulating that scrambling is uniformly a feature-mediated process, which is triggered by a scrambling feature Σ , as stated below.

¹¹As mentioned before, the status of scrambling in a given scrambling language is still a controversial matter; particularly, for Russian and German. However, as for other scrambling languages, there seems to be some consensus.

¹² Grewendorf and Sabel (1999), putting aside the cases of long-distance scrambling, argue that languages can employ either A-scrambling or A-bar scrambling, but not both. They consider German as an A-bar scrambling language and Japanese as an A-scrambling language. See detailed discussion on this approach in the following section.

(29) Scrambling as a feature-mediated process (Grewendorf and Sabel 1999:25)

Scrambling is a feature-mediated process driven by a scrambling feature Σ that is optionally realized with Agr-heads.

They assume that the scrambling feature Σ is optionally inserted in the Agr-head, and the optionality of realizing the scrambling feature in Agr-head is not subject to economy.

In this theory, the difference between A-scrambling and A-bar scrambling is attributed to different landing sites of scrambling; A-scrambling is a movement to Spec AgrP (L-related) positions, and A-bar scrambling is a movement to AgrP adjoined (broadly L-related) positions. Spec AgrP positions and AgrP adjoined positions are argued to have different syntactic properties. Spec AgrP positions bear A-properties, and can function as intermediate landing sites for long-distance scrambling, while AgrP adjoined positions bear A-bar properties, and are conceived as 'dead-ends' for every kind of movement.

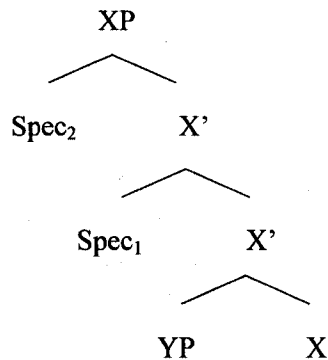
Grewendorf and Sabel argue that in order to allow movement to the Spec AgrP position in addition to subject movement to Spec Agr position, the agreement head must have multiple features to license multiple specifiers, in addition to having scrambling feature. This means that A-scrambling is available only in languages with multiple specifiers. Further, given that only the Spec AgrP position can serve as the intermediate landing site for movement, it also implies that long distance scrambling depends on the availability of multiple specifiers. The theory makes two fold claims, as stated below.

(30) Conditions for A-Scrambling (Grewendorf and Sabel 1999: 3)

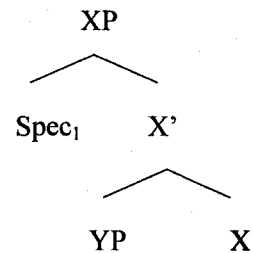
A scrambling language allows A-scrambling as well as scrambling out of finite clauses if and only if multiple Agr-specifiers are licensed in the language.

If the agreement head does not license multiple specifiers, scrambling can only target AgrP adjoined position (Spec Agr position is filled by the subject), forcing the scrambling element to land in the adjunction site. This yields A-bar properties of scrambling and also blocks long-distance scrambling. This is schematized below.

(31) a. A-Scrambling



b. A-bar Scrambling



In languages with multiple specifiers, (short) scrambling can proceed into Spec2 if the head X has the scrambling feature, as in (31a). Languages without multiple specifiers allow only adjunction to XP, as in (31b). In this system a given language cannot have both A-scrambling and A-bar scrambling at the same time.¹³ Depending on the availability of multiple specifiers, it will exhibit either A-scrambling or A-bar scrambling. Grewendorf and Sabel (1999) argue that scrambling in Japanese belongs to the first type, and scrambling in German belongs to the second type.

This approach has two implications for cross-linguistic variation: i) A-scrambling languages should be the subset of languages with Multiple Nom case marking (MNC), which is assumed to employ multiple specifiers.¹⁴ ii) Long-distance scrambling languages should be the

¹³ A-bar properties of pre-subject scrambling such as reconstruction effects are argued to be inconclusive. However, the issue seems to be controversial at this point.

¹⁴ Grewendorf and Sabel (1998) assume that an example of Multiple Nominative Case marking in (ia) requires multiple specifiers as shown in (ib).

(i) a. Mary-ga meri-ga kiltta.
 M-Nom hair-Nom long
 'Mary has long hair.'

subset of A-scrambling languages. According to Sugisaki's (2003) typology, the first implication seems to be borne out.

(32) Table 2. Cross-linguistic survey of scrambling and MNC

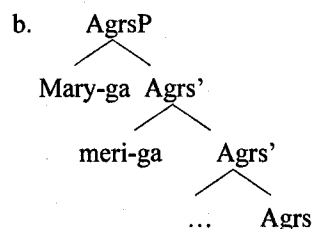
(Sugisaki: 2002)

	+MNC	-MNC
-Scrambling	Chinese, Modern Hebrew	English, French
A-Scrambling	Korean, Japanese, Modern Persian	
A-bar scrambling		Hindi, German, Turkish, Serbo-Croatian.

Sugisaki reports that for the small set of languages he surveyed, all languages which allow A-scrambling also exhibit MNC. However, as for the implicational relation between long-distance scrambling and A-scrambling, the prediction is not borne out. As pointed out in section 3.2, languages such as Turkish and Serbo-Croatian allow long-distance scrambling, but they do not exhibit A-scrambling properties, contrary to their predictions.

3.4.3. Case-related Approach: Bošković (2004, in press)

Bošković (2004, in press) takes scrambling as a unitary phenomenon, like the scrambling feature approach. Bošković's scrambling theory is partly based on his early work in Bošković and Takahashi (1998). They argue that scrambling as optional overt movement is not compatible with



Chomsky's (1994) conception of movement as a last resort operation. As a solution to the problem, Bošković and Takahashi explore a radically different approach to scrambling. They propose that "scrambled" phrases are base-generated in their surface positions, and undergo LF movement to their thematic positions under Last Resort. The Last Resort nature of LF lowering is ensured by theta-features. They assume that theta-roles are formal features which can be the driving force for movement (see Bošković 1994, Lasnik 1995, Hornstein 1999, for independent arguments for this approach). Let us take a concrete example to see how this mechanism works.

(33) John-ul_i [Mary-ga t_i Susi-ga mannatako] malhaesstta

J-Acc M-Nom S-Nom met-Comp said

'Mary said that Susi met John.'

Under this approach, the sentence (33) is derived in the following way. The embedded object *John-ul* is introduced in the matrix IP adjoined position by merge, and remains there for the PF component, but in LF, the object undergoes movement into the VP-complement position to receive a theta-role from the embedded verb. Otherwise *John-ul* would not receive a theta-role, and would remain uninterpretable at LF, causing the derivation to crash. Hence lowering of the object is forced by theta-feature checking, as an obligatory operation.

How does this theory derive the dual property of scrambling? Bošković and Takahashi rely on the possibility of reanalyzing the IP adjoined position as an additional specifier position. When the IP adjoined position can be reanalyzed as an additional [spec IP] at LF, it dispenses the need for lowering, giving rise to A-properties. Otherwise, we get A-bar properties. Let us see how A-property such as anaphor binding is accounted for in the base-generation analysis of scrambling.

(34=15b) kutul-ul_i selo-uy_i chinku-ka t kosohaysstta.

they-Acc each other-Gen friend-Nom sued.

‘Each other_i’s friends sued them_i.’

In (34), the object is base-generated in the IP adjoined position in overt syntax. In LF, the adjoined position is reanalyzed as an additional specifier, allowing the features (possibly theta-feature and Case feature) of the object to be checked after V-to-T movement, without lowering. Given that the object need not lower, it will be in a position to bind the anaphor. Since the derivation in question also depends on V-to-T movement (so that the object that does not lower can get a theta-role and case from the verb), it is available only in the case of short distance scrambling, not long-distance scrambling. When the reanalysis does not take place, the object will have to lower to the position where it can undergo necessary checking. Bošković and Takahashi (1998) conjecture that the possibility of reanalysis is parameterized; it is available only in languages with multiple specifiers, relating A-scrambling to the availability of multiple specifiers. This mechanism is basically similar to the assumption in Grewendorf and Sabel’s (1999) theory of A-scrambling. Both theories attribute the difference between A-scrambling and A-bar scrambling to the availability of multiple specifiers.

Having discussed the derivation of A and A-bar scrambling, let us turn to the question of the source of the cross-linguistic variation regarding scrambling. What allows languages to employ base-generation of an argument in this theory? Building on his scrambling analysis in Bošković and Takahashi (1998), Bošković (2004, in press) offers the following proposal.¹⁵ Bošković’s theory draws on two cross-linguistic facts concerning scrambling: lack of determiners and availability of overt case marker. He argues that scrambling languages do not have

¹⁵ Bošković (2004) actually hints at the analysis in a footnote. The analysis is further explored in Bošković (in press).

determiners (i.e. articles).¹⁶ His observation is from the contrast between Slavic languages and Romance languages. Most of Slavic languages do not have articles and allow scrambling, while modern Romance languages do have articles and do not allow scrambling. Japanese, Korean, Turkish, Hindi, and Walpiri, all scrambling languages also do not have articles.¹⁷ He considers those languages without determiners are NP languages, assuming that NPs in those languages are literally NPs.

He relates the NP vs DP distinction to another cross-linguistic fact relevant to scrambling: the availability of overt case marker. Bošković postulates a principle called ‘Argument Identification Requirement’, which states that an argument needs to be identified as such in overt syntax. He proposes that there are two ways to satisfy this requirement: overt case or merge in a theta-position. This means that argument NPs with overt case markers can merge in a non-theta position without violating the principle since they satisfy the principle via case markers. This gives rise to surface order of scrambling. However, argument DPs cannot merge in a non-theta position since they must merge in their theta-position in order to satisfy the Argument Identification Requirement. Bošković assumes that determiners impose merge in a theta-position. Since N, not D, is the actual source of case (in some languages D gets case via low level morphological spreading), only NPs but not DPs are case-marked in the syntax. Only NPs can then satisfy the Argument Identification Requirement via case marking. If a language has a determiner system, it always requires arguments to merge with the verb in a theta-position to satisfy the Argument Identification Requirement. Under this view, languages with determiners (i.e. articles) are not expected to allow scrambling regardless of the presence of overt case morphology.

¹⁶ Bošković thus argues that scrambling languages do not have DP/D. They may have elements corresponding to English *this*, *that*, but in several scrambling languages he examines they are adjectives, not Ds.

¹⁷ Recall that Bošković (in press) argues that German does not have scrambling. Grewendorf (in press) also indicates that what has been considered scrambling in German is actually either topicalization or focalization.

To summarize, all three analyses presented above achieve the merit of dispensing with optionality by postulating a morphological feature: EPP and Focus in Miyagawa (2003), or scrambling feature in Grewendorf and Sabel (1999), or by taking base-generation of an argument in non-theta positions as an option available in UG, when it does not violate any other principles in Bošković (2004, in press). The dual property of scrambling led in two different directions: one direction which takes scrambling as a non-unitary phenomenon and one direction which takes it as a unitary phenomenon. The EPP and Focus approach by Miyagawa (2003) belongs to the former, and the Scrambling feature approach by Grewendorf and Sabel (1999) and the Case-related approach by Bošković (2004, in press) belongs to the latter. All three approaches recognize A-scrambling and A-bar scrambling. The latter two approaches derive them from language specific phrase structure differences such as the availability of multiple specifiers, while the former postulates two distinct triggering features for each type of scrambling. However, unlike the EPP/Focus approach and the Case-related approach, the Scrambling feature approach assumes that two different types of scrambling exist only as a point of cross-linguistic variation, not as properties co-existent within a single scrambling language. Having discussed three major syntactic analyses of scrambling and their cross-linguistic predictions, we now consider acquisitional implications of each theory of scrambling.

3.5. Acquisitional Implications

In this section, we discuss implications of scrambling theories for three acquisitional issues: i) the acquisition order of A-scrambling vs A-bar scrambling, ii) acquisitional relation between A-scrambling and MNC, and iii) type of positive evidence required for learning scrambling.

All three scrambling theories recognize two types of scrambling. According to the EPP/Focus approach, a given language may have either type of scrambling depending on which feature is involved. In the Case-related approach, the availability of multiple specifier determines

which type of scrambling is allowed in a given scrambling language. Grewendorf and Sabel's (1999) theory is different from these two approaches in that they assume that a given scrambling language has only one type of scrambling, but never a mixture of the two. Hence, the theory does not make any language-internal predictions on the relative order of acquisition of the two types of scrambling. Unlike Grewendorf and Sabel's (1999) approach, the EPP/Focus approach and the Case-related approach raise a question concerning the possibility that the two types of scrambling may be acquired at a different point if a given language exhibits a mixture of A-scrambling and A-bar scrambling.

What are the acquisitional predictions? Which type of scrambling is likely to be acquired earlier than the other? The Case-related approach by Bošković (2004, in press) makes a clear prediction. In his theory of scrambling, A-scrambling and A-bar scrambling make implicational relations. Both types of scrambling require overt case marker and lack of articles,¹⁸ but one more condition is required to allow A-scrambling: multiple specifiers. This implicational relation makes a certain acquisitional prediction. It predicts that children will never acquire A-scrambling earlier than A-bar scrambling because A-scrambling requires the learning of the additional property. When both types of scrambling are tested, we do not expect to find children who pass A-scrambling to fail A-bar scrambling, but expect that some children who pass A-bar scrambling may fail A-scrambling, because A-scrambling requires an additional property.

The EPP/Focus approach seems to make the same prediction, but on different grounds. It is known in the acquisition research that children misinterpret passive sentences, considering them as if they are active sentences. Some researchers have attributed children's failure with the passive sentences to their delay in A-movement.¹⁹ Borer and Wexler (1987) propose that the

¹⁸ Bošković (2004, in press) leaves open the possibility that other factors, in addition to these two, may influence the availability of scrambling.

¹⁹ There have been counterarguments to this account. For instance, Grozinsky and Reinhart (1993), Snyder and Stromswold (1997) argue that children's problem with passive sentences is not with the A-movement property, but with theta-transmission involved in interpreting by-phrase. The issue is still unsettled.

machinery that allows the formation of A-chains is subject to maturation. If we adopt the ‘maturation of A-chains’ argument, we predict that A-bar scrambling is likely to be acquired earlier than A-scrambling because the EPP-driven scrambling belongs to A-movement. We expect that some children who pass A-bar scrambling still may fail A-scrambling, but not the other way around.

Sugisaki (2000) provides experimental data that shows that A-scrambling is indeed acquired later than A-bar scrambling. He conducted an experiment with Japanese monolingual children with age range between 3;8 and 4;6 on the following two types of ditransitive constructions: one with the direct object appearing to the front of the indirect object, as in (35a), and one with the indirect object scrambled to the sentence initial position, as in (35b).

(35) a. Satoshi-ga Pikachu-o_i okaasan-ni t_i misetayo.

S-Nom P-Acc mother-Dat showed

‘Satoshi showed Pikachu to his mother.’

b. Kasumi-ni_i Satoshi-ga t_i Pichu-o misetayo

K-Dat S-Nom P-Acc showed

‘Satoshi showed Pichu to Kasumi.’

The purpose of the experiment was to evaluate two different syntactic views on the structure of (35a). Miyagawa’s (2003) base-generation analysis holds that the order in (35a) is derived by base-generating the direct object preceding the indirect object. Hoji’s (1983) scrambling analysis holds that (35a) is a derived order as a result of A-scrambling of direct object crossing the indirect object.²⁰ This sentence was compared with (35b), in which the indirect object is scrambled to the

²⁰ It is a well-established fact that clause internal scrambling always exhibits A-movement properties (Saito 1985).

beginning of the sentence. Pre-subject scrambling such as (35b) is known to be ambiguous between A-scrambling and A-bar scrambling (recall the discussion in section 3.3).

In constructing acquisitional predictions, Sugisaki hypothesizes that the A-movement delay may carry over to the acquisition of A-scrambling, expecting that A-scrambling will be acquired later than A-bar scrambling due to the A-movement delay. If (35a) is derived by A-scrambling of direct object, and A-movement delay plays a role, children are expected to perform better with (35b), because (35b) can be analyzed as either A-scrambling or A-bar scrambling. It turns out that most children performed 90% correct with (35b), but failed 60% with (35a), as hypothesized. The result led Sugisaki to conclude that (35a) is indeed an instance of A-scrambling. The adult-like performance of (35b) was interpreted as a result of using A-bar scrambling strategy. This experimental data is consistent with both the EPP/Focus approach and the Case-related approach, but for different theoretical reasons. A-movement delay is the relevant factor for the EPP/Focus approach, while the implicational relation is the relevant factor for the Case-related approach.

Second acquisitional implication concerns the relation between A-scrambling and multiple Nominative case marking (MNC), as in (36), in which multiple elements appear with Nom case.

(36) Mary-ga Billy-ga chotta.

M-Nom B-Nom like

‘Mary likes Billy.’

Both Grewendorf and Sabel (1999) and Bošković (2004, in press) assume that A-scrambling is available only in languages with multiple specifiers. These two approaches make an acquisitional prediction that children will never acquire A-scrambling earlier than MNC because multiple specifiers are employed in MNC. Accordingly, when MNC and A-scrambling are tested, we

expect that there would be no children who pass the scrambling test, but fail the MNC test because A-scrambling requires scrambling feature, in addition to multiple specifiers.

Sugisaki (2003) conducts an experiment with Japanese learning children with the age range from 3;4 to 5;3 (mean age 4;6) to evaluate this prediction.²¹ Assuming with Grewendorf and Sabel (1999) that Japanese scrambling is only A-scrambling, he tested acquisition of MNC and scrambling. To test the acquisition of MNC, Sugisaki uses sentences which are made structurally ambiguous between MNC-reading and non-MNC reading. He checked whether the child accepts the MNC reading when the context is biased toward the MNC reading. As for the test for scrambling, he uses sentences with multiple scrambling to ensure that a property tested involves scrambling rather than topicalization, assuming that a single object scrambling may be interpreted by using topicalization. He reports that all children who passed Scrambling also passed MNC (See the footnote for results.).²² Based on this results, Sugisaki argues for Grewendorf and Sabel's (1999) scrambling feature theory of scrambling. Logically, children could use A-bar scrambling strategy to interpret multiple scrambling, but they did not, according to the result.²³ Sugisaki (2003) takes it as an indication that there is no parameter resetting from A-bar scrambling to A-scrambling.

Can the results be consistent with the Case-related approach, which also links the availability of A-scrambling to the availability of multiple specifiers? It seems not. In the Case-related approach, Japanese is considered to be a language with both A-scrambling and A-bar scrambling. This means that under the Case-related approach, the subjects in Sugisaki's study could have used A-bar scrambling as an alternative strategy to interpret multiple scrambling

²¹ The test method used in Sugisaki (2003) was truth-value verification (Crain & Thornton 1998).

²² The results of Sugisaki (2003) are shown in the following table.

(i) Performance on the MNC test and the MS test

MNC	MS test	
	Pass	Fail
Pass	8	2
Fail	0	6

²³ This fact may seem contradictory to the previous study by Sugisaki (2000).

sentences. Given that A-bar scrambling does not have an implicational relation with MNC, no contingency between passing scrambling and passing MNC is expected, contrary to the results reported in Sugisaki (2003).²⁴ In section 4, 5, I will suggest a solution to this potential problem for the Case-related approach.

Finally, let us consider what each scrambling theory predicts concerning the issue of the positive evidence required for scrambling. As for the Scrambling feature approach, given that the scrambling feature Σ does not have any morphological reflex, it seems that scrambling utterances themselves are the only evidence for both types of scrambling. On the other hand, Bošković's Case-related approach allows the possibility that the child may figure out the availability of scrambling by recognizing the overt case marker and lack of articles, even without hearing actual scrambling sentences. A-bar scrambling may be signaled by overt case marker and lack of articles, but A-scrambling would require one additional property, MNC sentences, besides overt case marker. In the EPP/Focus approach, A-scrambling requires V-to-T raising and overt case morphology. It seems that the theory predicts that the child may not need to hear scrambling utterances to learn the availability of A-scrambling. As for A-bar scrambling, utterances with preposed focused element will indicate the availability of A-bar scrambling.

The previous experimental studies of scrambling concern the acquisition order with A-scrambling vs A-bar scrambling and the possible relation between A-scrambling and MNC. We have seen that the acquisitional results in those studies could potentially support multiple theories, not just a single theory of scrambling. Further, they did not address the most crucial issue, a learnability question: what kind of linguistic evidence does the child use to determine whether their language allows scrambling or not? We have drawn implications of each theory of scrambling for the positive evidence required for the learning of scrambling. In the following chapter, we subject them under a further investigation.

²⁴ A possible future research may be testing uncontroversially A-scrambling sentences along with MNC sentences, to verify the implicational relation between MNC and A-scrambling.

Chapter 4. Learnability Problem in Scrambling

4.0. Introduction

In chapter 3, we reviewed major grammatical properties of scrambling. We observed: i) that scrambling is an optional operation in the sense that canonical order is always available as an alternative, and ii) that scrambling exhibits both A and A-bar movement properties, surfacing as two distinct types, A-scrambling and A-bar scrambling. Three syntactic theories of scrambling were discussed: the EPP/Focus theory (Miyagawa 2003), the Scrambling feature theory (Gewendorf and Sabel 1998), and the Case-related theory (Bošković 2004, in press).

The EPP/Focus theory by Miyagawa (2003) takes scrambling as an operation involving two distinct types of features: EPP feature and Focus feature. A-scrambling is analyzed as a movement driven by EPP feature, and A-bar scrambling as a movement driven by Focus feature. The theory argues that two conditions are required to allow A-scrambling: V-to-T movement and overt case morphology. V-to-T movement extends the checking domain of T, making the object equally local to T as the subject. Overt Acc case enables the object to be licensed by T, allowing it to move to spec TP. A-scrambling is the instance of the object moving to spec TP for EPP-feature checking, with the subject staying in the base-generated position.

The Scrambling feature theory by Gewendorf and Sabel (1998) postulates scrambling feature Σ as a driving force for scrambling. The distinction between A-scrambling and A-bar scrambling is attributed to the type of position that the scrambled element lands in. If it lands in Spec AgrP position, A-scrambling properties result. If it lands in the AgrP adjoined position, A-bar scrambling properties result. The availability of Spec AgrP position as the landing site of scrambling is conceived as a parameterized property on the Agreement head. Gewendorf and Sabel (1998) argue that the Spec AgrP position as the landing site of scrambling is available only

in languages whose Agr head can check multiple sets of case features, relating A-scrambling to the availability of multiple specifiers.

Finally, the Case-related theory by Bošković (2004, in press) argues that scrambling is lexical insertion of arguments in non-theta positions, as an instance of pure-merge. Assuming that NP arguments in languages without determiners (i.e. articles) are really NPs (lexical projection), Bošković argues that they can merge in a non-theta position. He postulates a principle concerning licensing of arguments, which states that an argument must be identified as such in the overt syntax through either overt case or merge in the theta-position. This implies that only argument NPs with overt case marker can merge in a non-theta position without violating the principle since the overt case marker satisfies the principle. Argument DPs must always merge in the theta-position in the overt syntax since they can be identified as arguments only by merge with the verb, because the NP, which is the source of case, is not the actual argument in DP languages, the NP being dominated by DP. Under this theory, two conditions are required for scrambling: overt case and lack of determiners.

In this chapter we investigate a learnability question: how does the child acquire scrambling? Particularly, what kind of linguistic evidence do children rely on to acquire scrambling? A simple assumption is that they would learn it by hearing scrambling utterances, as exemplified in (1).

(1) John-ul_i yojatul-i t_i milesstta.

J-Acc girls-Nom pushed

‘The girls pushed John.’

This seems to be a fairly reasonable idea; after all, scrambling is a language particular property, and it seems natural to assume that there should be enough scrambling utterances available in the input. However, we can think of another acquisition scenario, postulating less direct relation

between scrambling and the relevant positive evidence. Suppose that there is a property X, which indicates the availability of scrambling somehow. The child may detect the property, and use it to acquire scrambling. We call this property X a cue. The first acquisition scenario seems very intuitive. It simply assumes that the positive evidence for scrambling is the very sentences which illustrate the scrambling properties, and that the scrambling sentences are reliably available in the input. The second scenario seems less intuitive; if there is enough direct positive evidence for scrambling utterances, why should we consider more abstract kind of positive evidence additionally?

However, when we try to take it into account that some scrambling languages do not have overt topic or focus markers, things do not seem so straightforward as they look. In those scrambling languages without overt topic or focus marker, the surface order such as (1) would not be enough evidence for scrambling because they can be ambiguous between scrambling and topicalization or focalization. This implies that other evidence which can reveal unique property of scrambling would be required to indicate the availability of scrambling: for instance, samples of multiple scrambling or scrambling involving anaphor binding, etc. This brings in an issue of whether such positive evidence would be reliably available in the input or not, adding a bit of complication to the first acquisition scenario. On the contrary, the second acquisition scenario does not require any further complication of this kind. Since it is based on the idea that there is a distinct cue for scrambling, it automatically implies that such cue be available as long as a given language allows scrambling. Further, since cue is not the direct scrambling sentences, it allows the possibility that the child may learn scrambling without hearing the actual scrambling utterances from the input. Now, both acquisition scenarios seem to have merits.

To explore the above two acquisition scenarios, we examine the child speech and parental input of scrambling from longitudinal production data. We find a close resemblance between the child speech and the adult input patterns. Children rarely produced even simple scrambling utterances such as the ones in (1), let alone multiple scrambling or other types of scrambling.

Similarly, scrambling utterances in the input were extremely few. Determining that children cannot acquire scrambling by hearing scrambling utterances given the impoverished input, we consider the second acquisition scenario which involves learning of scrambling by a cue.

Based on some cross-linguistic facts and a phenomenon of ‘case dropping constraint’, we hypothesize that Acc case marker can be a potential cue for scrambling. We conduct an experimental study to test whether there is a systematic association between acquiring scrambling and acquiring overt Acc case marker. The results show a significant contingency between passing Acc case test and passing scrambling test, suggesting that the child makes use of Acc case marker as a cue for scrambling. These experimental results have implications for the syntactic theory of scrambling and the nature of triggering experience in language acquisition.

The chapter is organized in the following way. Section 4.1 examines the longitudinal data of the child speech and the adult input for scrambling utterances. Section 4.2 reviews the early acquisition studies of scrambling. Section 4.3 considers arguments for Acc case as a cue for scrambling. Section 4.4 explores experimentally the acquisitional predictions for the possible relation between Acc case and scrambling. Section 4.5 offers some speculation on conflicting results between Sugisaki (2000) and the present study. Chapter 4.6 and 4.7 discuss implications of the results for the theory of scrambling and for the theory of learnability. Section 4.8 is the conclusion of the chapter.

4.1. A Learnability Puzzle

In this section, we examine longitudinal production data of scrambling in both child speech and adult input to check the first acquisition scenario. We first look at the child speech data, and then compare it with the parental input. We did not find the data for a direct figure of scrambling utterances, but we did find some relevant information. Table 1 below shows word order patterns in child speech.

(2) Table 1. Frequency of word order patterns in the child speech
(Kim 1985: 350)¹

Word Order	Utterance n/%		
Child (age range)	Alicia(2;3-2;9)	Paul(2;7-2;11)	Anne(2;10-3;4)
SV	364/38.6%	523/39.1	698/37.3
VS	14/1.5%	17/1.3%	22/1.2%
OV	288/30.5	388/29.0%	591/31.6%
VO	10/1.1%	10/0.7%	194/10.4%
SOV	200/21.2%	334/25%	473/25.4%
OSV	12/1.3%	16/1.2%	39/2.1%
SVO	16/1.7%	11/0.8%	5/0.3%
OVS	38/4.1%	37/2.8%	34/1.9
Total	943/100%	1,337/100%	1,869/100%

The frequency of OSV gives us some idea of the approximate proportion of scrambling. The frequency of OSV may include instances such as object topicalization, or object focalization, in addition to object scrambling. This means that the actual proportion of scrambling would be less than this figure, if any. Logically, it is also possible that no instance of OSV is a scrambling utterance. In any case, it seems very likely that the frequency of scrambling would be extremely low, given the low frequency of OSV.

How should we interpret this low frequency of scrambling utterances in child speech? Can we conclude that children did not have the knowledge of scrambling, and as a result, they could not produce scrambling sentences? This would be easily confirmed if the input is abundant, and children still did not produce scrambling sentences. Then, we can be sure that the near absence of the data in the child speech has to do with the lack of relevant grammatical property. However, if the input is also very rare somehow, we cannot make any conclusions. There is the possibility that children may have knowledge of scrambling, but just they did not produce scrambling sentences, attempting to match the adult input pattern. In that case, the absence of

¹ This is based on Cho's (1982) data.

scrambling in the child speech may be only apparent. For this reason, we need to check input patterns of scrambling. Table 2 below shows the frequency of word order patterns in child-directed speech by a mother. The data is arranged by different age points of the child.

(3) Table 2. Frequency of word order patterns in MJ's mother

(Chung 1994: 114)

Word order	Age of the child			
	2;0	2;3	2;6	2;9
SOV	24/15%	28/12%	15/9%	19/12%
OSV	2 /1%	5 /2%	0	1/1%
SV <i>pro</i>	14/9%	9/4%	2/1%	4/2%
<i>Pro</i> OV	42/26%	28/12%	36/21%	25/15%
SV	76/47%	158/68%	112/67%	113/69%
VS	4/2%	2/1%	0	2/1%
Total	162	230	165	164

Surprisingly, very much like the pattern found in the child speech, the frequency of OSV is extremely low throughout the corpus: 1%, 2%, and 0. 1% (out of verb containing utterances). If we take three term utterances as the relevant denominator, the rate goes up to 7%. But again, recall that this figure still will not represent the actual number of scrambling utterances. In order to find out the frequency of scrambling, we need to locate OSV with the objects appearing with Acc, not the topic or focus marker. Cho's (1982) study provides us with the relevant data. That is shown in Table 3.

(4) Table 3. Frequency of Occurrence of the Acc marker in Mother's speech
(Cho 1982: 83)²

	Alicia's mother		Paul's mother		Anne's mother	
	+ case	- case	+case	- case	+ case	-case
OV	34	228	62	313	40	483
VO	4	5	2	7	0	5
SOV	2	87	3	155	4	160
OSV	2	8	3	11	0	32
OVS	1	17	3	15	1	14
SVO	2	14	2	9	1	3
VSO	0	0	2	1	1	1
VOS	1	0	0	0	0	0
Total	45	359	74	511	47	698

The above table shows the frequency of occurrence of Acc marker out of potential contexts from three mothers' speech. OSV utterances with case markers are analyzed as scrambling utterances. OSV without case marker are instances in which the object is topicalized, but the topic marker is (grammatically) dropped (see section 4.3 for the discussion on the obligatory status of Acc marker for scrambled objects.). OSV utterances with case markers are scrambling utterances. The frequency is still extremely low: 2, 3, 0. This number corresponds to 0%-1.5% out of three term utterances, and 0.003% out of all object containing utterances.³ We examined Jiwon's longitudinal corpus from CHILDES data base (MacWhiney 2005) to verify the low frequency of scrambling reported above. The results confirmed Cho's finding, as shown in the following table.

² The subjects in Cho (1982)'s study are the children of Korean parents living in Canada. She mentions that the subjects were carefully selected on the following basis.

- i. Children never attended day care centers where English is spoken.
- ii. Mothers spoke only Korean at home and they were not employed.
- iii. Children had no English speaking friends.

³ The frequency of scrambling seems compatible with the frequency of OVS or SVO, which are only marginally acceptable. Korean being strictly head-final, these non-head-final utterances are considered to be almost ungrammatical, unless a strong stress is applied. The ungrammatical status of OVS, SVO or VOS is the same regardless of the presence of the Acc marker. These utterances may be very likely to be cases of performance error, or after thought insertion.

(5) Table 4. Patterns of Object marking in parental input

Word order	Jiwon (2;0)	
	+ case	- case
(S)OV	109	440
OSV	0	1

Out of over 500 utterances with overt objects, none was a scrambling utterance.

The rare input of scrambling is rather surprising, given that scrambling does not involve any complications such as depths of embedding. Why are scrambling utterances so few? There seem to be a couple of possible factors that may be involved in reducing the number of scrambling utterances. The first is the optionality. As discussed in the previous chapter, scrambled word order is a variant of canonical word order. There is no particular context where scrambled order is required, as opposed to canonical order. Given this, adults simply may have chosen to use canonical word order just for a preference. The second possible factor is the productive use of null argument *pro*. A large proportion of utterances (about 70%) are instances of two term utterances, whereas three terms are needed to detect scrambling. This may have reduced the chance of using scrambling. This extreme paucity of scrambling in the input leaves the nature of the lack of scrambling in the child speech unclear. Given that the adult input is also rare as the child speech, we cannot be sure that the absence of scrambling in the child speech is due to the lack of the relevant knowledge of scrambling. It could be that children simply preferred not to use scrambling, just like their adults.

Considering the above child speech data, one may assume that the child would acquire scrambling quite late, but some recent experimental studies show that that is not the case. Otsu (1994) showed that Japanese children acquire scrambling as early as the age of 3, demonstrating that scrambling is acquired along with other grammatical properties, crucially not particularly late. Murasugi (2002) confirmed this finding.

Positive evidence for scrambling is extremely rare, but nonetheless, children seem to be acquiring scrambling successfully, as evidenced in the experimental studies. This poverty of stimulus situation raises a learnability question. How do children acquire scrambling if they do not hear enough scrambling utterances? How can we account for the gap between the input and the output? In dealing with the poverty of stimulus situation, it is customary to attribute the relevant part of knowledge to innate components. But, what we are dealing with here is the acquisition of a parametric property which has to be learned from experience. As a possible solution to this problem, we turn to the second acquisition scenario. There may be robust input of some property X, which indicates the availability of scrambling. This property may serve as a cue for the acquisition of scrambling. Then, the question is what that cue could be. Before we explore this question further, we review early acquisition studies of scrambling in the following section.

4.2. Review of Early Acquisition Studies of Scrambling

With Chomsky's (1959) introduction of linguistic universals, the acquisition of word order came to be the focal point for investigating the role of universals in language acquisition. Slobin (1965) observed that English learning children as well as Russian children initially used fixed word order. Pointing out that Russian children went through the fixed word order stage, although Russian allows variable word orders, Slobin (1965:134) argues that children are preprogrammed to assume language to be ordered, and to adopt a basic order as a first guess, and later learn that the basic order can be changed.

However, in his later experimental study, Slobin (1982) observes that children learning Turkish, which he identifies as a minimally ordered inflectional language, did not have trouble comprehending non-canonical order from early on, showing that they start with variable word orders instead of trying fixed word order at a prior stage. From this, Slobin hypothesizes that languages differ as to how grammatical relations are encoded: by word order or by case inflection.

The underlying assumption was that case inflection indicates grammatical relations, and as a result the order of elements can be shuffled around; no particular ordering of elements is required.

4

Such assumption can be traced back to an old cross-linguistic generalization in traditional grammar studies. It has been often observed that languages with rich case morphology seem to allow flexible word order.⁵ For instance, Sapir (1921) recognized a tendency in language change that seems to hold between the loss of case distinctions and the loss of free word order. Jespersen (1922:361) made a similar observation:

This, then, is the conclusion I arrive at, that as simplification of grammatical structure, abolition of case distinction, and so forth, always go hand in hand with the development of a fixed word order, this cannot be accidental, but there must exist a relation and effect between the two phenomena.

Hawkins (1986:40) and Blake (1994:15) also mention the same generalization.

With Slobin's pioneering work on acquisition studies on word order, various languages have been studied to evaluate this argument. For instance, Chung (1994) tested Korean children's comprehension on sentences with canonical order and sentences with scrambled order, and found that children in group II (age range: 3;3-3;5) performed 80% correct on canonical order, but only 50 % correct on scrambled order. Studies with Japanese children report the same fact (Miyahara

⁴ According to this hypothesis, languages such as English express grammatical relations by fixed word order, while languages such as Turkish rely on inflection because word order is so variable.

⁵ There have also been some studies which relate variation in basic word order across languages to overt morphological case. For instance, Greenberg (1966) observes the following universal tendency: "if in a language the verb follows both the nominal subject and the nominal object as the dominant word order, the language almost always has a case system." There is at least one language which does not seem to fall into this general tendency. Navajo is SOV word order language, but does not have any overt case marker. However, by and large there seems to be a strong tendency of SOV languages to have overt case morphology. Blake (1994) reports the following results of his survey;

SVO	[+case] 9	SOV	[+case] 34
	[-case] 26		[-case] 7

1971, Hayashibe 1975, Hakuta 1982, and Clancy (1986) for Japanese. Such studies analyzed children's failure to comprehend the scrambled word order as an indication that those children preferred word order cue to inflection cue. The successful performance of older children was analyzed as the indication that inflection cue is acquired later than word order cue.

According to Slobin's hypothesis, the results on Japanese and Korean are unexpected, given that both are inflectional language in his classification. Chung (1994) offers an account using competition model of grammar (Bates and MacWhiney 1982, MacWhiney et al. 1984, Bates and MacWhiney 1987). She argues that the word order in Korean is a more reliable cue than case marker, and hence the child relies on word order prior to case markers.⁶ Along the same lines, Clancy (1985) conjectures that the different results between Turkish and Japanese may be due to the fact that in Turkish the word order may be freer than in Japanese, and that inflection cannot be omitted.

The above early acquisition studies of scrambling considered case inflection as a key to acquire scrambling. They assumed that children acquire scrambling when they learn case inflection. Relating the acquisition of scrambling to case inflection seems to allow the possibility that children can learn scrambling without hearing scrambling sentences. When we consider the finding that children are not likely to acquire scrambling by hearing scrambling utterances, this early approach to acquisition of scrambling seems to have some merit.

However, such reasoning in relating case inflection to free word order in terms of identifying grammatical roles encounters an immediate problem. If identifying grammatical relation is what relates inflection to scrambling, we predict that the presence of either overt Nom or overt Acc would be enough to allow scrambling. For instance, if the subject is marked with overt Nom, it should allow the bare object to be scrambled since we can deduce the role of the object by the process of elimination. We will see that this is not the case (see example (7) in

⁶ Chung (1994) calculated overall cue validity of word order and case marker and compared them. Cue validity of word order was .71, while cue validity of case was .32.

section 4.3). Further, the idea of analyzing languages according to two different ways of encoding grammatical relations as to whether grammatical relations are encoded in inflection or word order has been abandoned as detailed syntactic studies have revealed more findings on the properties of non-configurational languages. It has been shown that grammatical relations are structurally determined also in non-configurational (i.e. free word order) languages and hence all languages have hierarchical structures.

The new perspective on non-configurational languages has led to a different way of analyzing scrambling. Scrambling is considered as a distinct syntactic phenomenon rather than a result of not being configurational. In this approach, the facts from early acquisitional studies of scrambling reported above take on a different interpretation. The failure of children to interpret the non-canonical word order is attributed to the lack of knowledge of scrambling, rather than to the incomplete knowledge of grammatical relations per se. Similarly, their ability to comprehend non-canonical word orders is attributed to their knowledge of a distinct syntactic operation, scrambling. If acquisition of scrambling is not about learning of a mode of encoding grammatical relations used in the target language, what is it that children have to learn?

4.3. Acc case Marker as a Cue of Scrambling

In the pervious section we have pointed out that the underlying assumption of the early acquisitional study in relating scrambling to overt case morphology incorrectly predicts that the presence of overt Nom case would be enough to allow scrambling. In this section, we argue that what is crucial in allowing scrambling is the presence Acc case marker, not the presence of any case marker. Cross-linguistic typology on the presence of Acc case marker and scrambling supports this argument. We find that all scrambling languages that have been identified have an overt Acc case marker.

(6) Table 5. Typology of scrambling and Acc marker

		<u>Scrambling</u>	
		Yes	No
<u>Acc</u>	Yes	Korean, Japanese, Chukchi, Persian, Hindi, Turkish, Russian, Serbo- Croatian, Tamil, Malayam...	Icelandic...
	No	*	English, Navajo, Swahili, Chinese, French, Thai

As shown in the table, not all languages with Acc case marker allow scrambling, but all scrambling languages have the Acc case marker. This cross-linguistic pattern seems to show that overt Acc case may indeed be related to the availability of scrambling.

A more convincing argument for the link between scrambling and Acc case is from the interaction between the presence of overt Acc case and the availability of scrambling found within some scrambling languages. In Korean, Acc case can be freely dropped when the object is in canonical position, but when the object is scrambled, the optionality of Acc case disappears, as shown in (7b).

- (7) a. Mary-ga John-(ul) milesstta.
M-Nom J-Acc pushed.
'Mary pushed John.'
- b. John-*(ul) Mary-ga t milesstta.
John-Acc M-Nom pushed
'Mary pushed John.'

Example (7a) shows that Korean allows Acc marker to drop. However, this option is no longer available when the object is scrambled, as shown in (7b), unless the object has the status of a topic. The bare object can appear in pre-subject position, but in that case, it is interpreted as a topic. In other words, the fronted objects without any marker are instances of topic marker drop (See Saito 1985 and Kuno 1973), not the instance of scrambling. The following set of data confirms this point.

(8) a. *nwuku-nun Mary-ga t mannassni?

who-Top M-Nom met

‘Who did Mary meet?’

b. nwuku-ul Mary-ga t mannassni?

Who-Acc Mary-Nom t met

c. *nwuku Mary-ga t mannassni?

Who Mary-Nom met

d. John-nun yojatul-i t milesstta

John-Top girls-Nom pushed

‘As for John, girls pushed him.’

e. John yojatul-i milesstta.

John girls-Nom pushed

Example (8a) shows that the topic marker can not appear on wh-words. (8b), which is the grammatical counterpart of (8a), shows that wh-phrase can scramble. Now, notice that when the Acc marker is dropped, as in (8c), the sentence becomes ungrammatical, demonstrating that scrambling requires Acc marker to be present. This means that example (8e) is the case in which the object drops the topic marker, not the Acc marker.

It has been known that some scrambling languages display a correlation between the presence of Acc case marker and semantic interpretation of the object. In those languages the Acc case marker appears only on objects which have ‘specificity’ interpretation. Most interestingly, scrambling is confined only to those specific objects. Turkish is one of such language, as demonstrated below (Enç 1991, Kornfilt 1994, Sener 2000, Gürel 2000).

(9) a. Elif elma-yi yedi

Elif apple-Acc ate

‘Elif ate the apple.’

b. Elif elma yedi

Elif apple ate

‘Elif ate apples.’

c. Elma-yi Elif yedi

apple-Acc Elif ate

d. *Elma Elif yedi

apple Elif ate

The contrast between (9a) and (9b) shows that the presence of Acc case has the effect of specific interpretation. Scrambling of the specific object is allowed, as in (9c); however, scrambling of the non-specific object induces ungrammaticality, as in (9d). The same fact is observed in Persian (Karimi 2003) and Tamil (Sarma 1999).⁷ My personal Tamil informant provided the following data in (10).

⁷ Even though the generalization holds, there are some language specific facts regarding the way the overt case marker interacts with the specificity of the object. For instance, Karimi (2003) points out that Persian has a small set of verbs in which the generalization breaks down. She shows that when those verbs are used, the objects take overt case marker regardless of specificity. In the case of Tamil, the generalization holds only for nouns that are considered [-rational], as shown below (Sarma 1999).

(10) a. John apple-ai saptan.

John apple-Acc ate

‘John ate the apple.’

b. John apple saptan

John apple ate

‘John ate apples.’

c. apple-ai John saptan.

apple-Acc John ate

d. *apple John saptan.

apple John saptan

The object with specific interpretation appears with Acc case, as in (10a), and it can undergo scrambling, as in (10c). However, the object without such semantic interpretation appears as bare object, as in (10b), and consequently scrambling is blocked, as shown in (10d).

Considering these data, one may wonder whether scrambling is determined by a semantic factor of specificity, rather than overt case marker itself. If that is the case, specificity is very likely to be a cue for scrambling. However, this line of reasoning does not seem to go through. Consider the following Korean data.

(11) a. Mary-ga cha-(lul) saassta

M-Nom car-Acc bought

‘Mary bought a car.’

b. Mary-ga ku cha-(lul) kociessta

M-Nom that car-Acc fixed

‘Mary fixed that car’

c. ca-*(lul) Mary-ga saassta

car-Acc M-Nom bought
 d. ku ca-*(lul) Mary-ga kociesstta
 that car-Acc M-Nom fixed

(11a/b) shows that regardless of the presence or the absence of overt Case marker, the object can be interpreted as either definite or specific object. Furthermore, non-specific object and specific object both can be scrambled, as shown in (11c) and (11d), as long as they contain overt case marker. This suggests that overt case marker in Korean does not encode specificity. This shows that specificity is not a grammatical basis for scrambling. This means that overt Acc case is indeed the relevant factor in allowing scrambling, not the semantic factor such as specificity.

Now, if overt Acc marker is indeed the relevant factor in allowing scrambling, is it plausible that Acc case is reliably available in the input as a cue for the positive evidence of scrambling? Do Korean children receive enough input of Acc case? It seems that they do. We have seen in the previous chapter that children receive Acc case about 20% to 30% of the time out of potential contexts, which is definitely higher than the rate of scrambling utterances and sufficient for learnability.

The cross-linguistic facts on scrambling and overt Acc case, the obligatory status of Acc case for the scrambled element within scrambling languages, and the availability of Acc in the input all suggest that acquiring scrambling by recognizing Acc as a cue for scrambling is quite plausible. In fact, there have been some anecdotal remarks in the previous spontaneous speech studies which alluded to a relation between Acc case and scrambling. For example, Chung (1994) made an observational remark that the Korean children began to scramble the word order after they had learned the accusative marker. However she offered a 'functional' account for the observation, proposing that the scrambling emerges as a response to their increasing pragmatic necessities. Chung suggests that children learn OSV order from topicalization and later produce the accusative marker in the construction, which results in scrambling. However, she does not

discuss the nature of the pragmatic necessity involved in scrambling. She seems to assume that scrambling is not so different from topicalization in that respect. This line of reasoning is problematic because it cannot explain the differences between topicalization and scrambling; for instance, the possibility of displacement of multiple elements.⁸

4.4. Experiment

4.4.1. Experimental Design

In order to test the link between Acc case and scrambling, we carry out an experimental study. We investigate whether there is any systematic relation between the acquisition of Acc case marker and the acquisition of scrambling. The experiment is to determine whether the acquisition of overt case morphology is required for the acquisition of scrambling or not. We construct the following two hypotheses.

(12) a. *Null Hypothesis*: There is no systematic association between scrambling and Acc case marker.

b. *Research Hypothesis*: Children will pass scrambling test only if they pass Acc test.

The null hypothesis predicts that there is no correlation between the acquisition of scrambling and the acquisition of Acc case marker, if case morphology is not a requirement for scrambling. However, if Acc case morphology is required to acquire scrambling, we expect that children will pass scrambling test only when they pass Acc test.

We use a comprehension task as the method of experiment, considering that both scrambling and the Acc case marker being optional, there would be no way to force the child to

⁸ A similar observation has been made concerning Japanese. Miyahara (1973) also reported that his Japanese subject started to use various word orders after he produced the Acc case marker '-o'.

produce them. To test Acc case, we make use of null argument *pro* subject, which is independently available in Korean. Since the subject is replaced by *pro*, the child has to know Acc case in order to comprehend the sentence.⁹

The following four types of test items were constructed.¹⁰

- (13) a. *pro* NP-acc V (Acc test)
 b. NP-nom *pro* V (Nom test : Control for Acc test)
 c. NP-acc NP-nom V (Scrambling test)
 d. NP-nom NP-acc V (Canonical order test: Control for Scrambling test)

The test items for the knowledge of Acc case marker take the form of (13a), in which the subject is realized as *pro*.¹¹ Item (13b), in which the object is realized as *pro* is a control item for (13a). The only difference between (13a) and (13b) being the case marker on the overt argument, the child should know each case marker in order to comprehend each type of sentence. So, the comparison between (13a) and (13b) will allow us to determine whether any failure in (13a) is due to lack of knowledge of Acc case or *pro*. Kim (2000) reports that Korean children start by producing overt subject and object 0% of the time (unlike older children or adults), but reach the adult rate between the age range of 1;10-2;2 for subject and of 1;7-2;0 for object. We can interpret this as indicating that they have knowledge about the status of subject/object *pro* by this time. The subjects in our study are over this age. (13c), in which the object is fronted to the sentence initial position, is the test item for the test of scrambling. (13d) with canonical order, is

⁹ Grammaticality Judgment task (e.g. Hiramatsu & Lillo-Martin 2001) could be ideal to test the knowledge of case markers. However, considering the relative young age of the subjects (from 2;2 to 3;2), we excluded the option.

¹⁰ See Appendix to this chapter for all items used in the experiment.

¹¹ As we will see in section 4.4.2, none of the predicates used for the Acc test are stative predicates, which would require Nom objects. They are all verbs that require Acc case for the object. I thank William Snyder for this comment.

used on a control item for (13c). Since both (13c) and (13d) have overt arguments with case morphology, differing only in word order, the performance on (13d) will help us to determine whether any failure on scrambling items is due to interpreting a sentence with two terms, or to the lack of knowledge about scrambling.

In each test, the test item is preceded by an additional sentence which provides background information. This is a necessary ingredient for testing (13a) and (13b) since they contain *pro*, which needs to be identified from the context. To make the whole test uniform, this step is taken for the other two test items (13c) and (13d). The above 4 test items allow us to check possible interaction between a particular case and scrambling. According to the traditional assumption on case inflection and scrambling, it is expected that the child may learn scrambling when they learn Nom case marker since that is enough to differentiate grammatical relations. On the other hand, if our research hypothesis is on the right track, we should expect to find a significant association between Acc case marker and scrambling.

4.4.2. Method

A picture identification task was chosen for the experimental method since children of relatively young age were targeted. The task involves having the child shown two different pictures, one on each side. The child is asked to point to the one which matches a sentence provided by the experimenter. The experiment started by familiarizing the child with the animals which show up in the pictures. The experimenter first approaches the child by asking what animal she likes, and shows the one that she answers, then moves to the other animals.¹² When the experimenter is confident that the child knows all the animals that appear in the pictures, she says to the child that they will play a guessing game, in which the task is disguised.

Before the actual test items are presented, a training test is given to ensure that the child understands the task. The pictures for the training test involve single animals doing different

¹² Most of the animals that appeared in the experiment were already familiar to the subjects.

actions. So, the sentences presented in the training test all involve intransitive verbs. For instance, a picture on left side has a turtle swimming in the water, and a picture on the right side has a turtle sleeping on the rock. The following sentence is presented before opening the picture book; ‘A turtle is sleeping on the rock! Which picture is it?’ Then, the experimenter opens the book, and the child is expected to point to the picture on the right side. The pictures were arranged in such a way that the matching picture appears sometimes on the left side and sometimes on the right side. When the child fails more than one training test item, she is excluded from the experiment.

The actual test items are different from training items in that the pictures involve two animals, which are involved in the same action, but in a reversed way. To take an example from the scrambling test, in one picture, a tiger chases a monkey, and in the other picture, a monkey chases a tiger. Before the two pictures are shown to the child, the following test item is presented.

(14) Sample for Scrambling test

a. supsoke horangi-ga issess nunte, ku horangi-ul wonsungi-ga chottagaeyo.

forest-in tiger-Nom was then that tiger-Acc monkey-Nom chased

There was a tiger in the forest. Then, a monkey chased that tiger.

Background

stimulus

Each test item consists of two parts; one part is the background information, and then the stimulus part, as indicated in (14). The items for the other tests take the same format, as illustrated below.

(15) Sample for Nom test

kongwon-e ttokki-ga issess nunte, kubuki-ga *pro* milesseyo

park-in rabbit-Nom be then, turtle-Nom pushed

‘There was a rabbit in the park, then a turtle pushed him.’

(16) Sample for Acc test

kongwon-e koyangi-ga issess nunte, *pro* ttokki-ul milesseyo

park-in cat-Nom be then, rabbit-Acc pushed.

‘There was a cat in the park, then he pushed a rabbit.’

(17) Sample for Canonical order test

supsoke-e taramci-ga issess nunte, ku daramchi-ga kangaci-lul coagasseyo

forest-in squirrel-Nom be then, that squirrel-Nom dog-Acc chased

‘There was a squirrel in the forest, then the squirrel chased a dog.’

There were five tokens for each type of test. Since there are four sentence types (Acc test, Nom test, Scrambling test, and Canonical order test), 20 trials are provided in total (See Appendix). Three types of verbs are used an equal number of times: pushing, chasing, and biting.

4.4.3. Subjects

The subjects who participated in the experiment are Korean monolingual children living in Seoul. They were recruited from several different day-care centers. As for the age range of the subjects, we assumed that children who are expected to be in the process of acquiring Acc case marker and scrambling should be included. To find out the appropriate age range, we relied on the reports from spontaneous speech data in previous studies. Spontaneous speech studies (Chung 1994 and Kim 1991 for Korean, Matsuoka 1999 for Japanese) reported that children acquire Nom between 1;9 and 2;2 and Acc between 1;11 and 2;5. Reports about the emergence of scrambling are less clear. Chung (1994) reports that scrambling utterances appear after 2;7 in spontaneous speech data, but with very few numbers. Since there is no available information about the point of productive use of scrambling in spontaneous speech data, we considered the report from the previous experimental studies (Otsu 1994 and Murasugi 2002), which showed that children

acquire scrambling as early as before the age of 3. Taking these facts together, we figured that it is reasonable to target children of the age range from 2;2 to 3;2.

4.4.4. Prediction

The hypotheses we set up in (12) give us three different patterns of distribution for logically possible responses. First, if children acquire scrambling independently of Acc case morphology, as in the null hypothesis, we expect the distribution in (18).

(18) Distribution of responses under Null hypothesis

		<u>Scrambling</u>	
		Pass	Fail
<u>ACC</u>	Pass	√	√
	Fail	√	√

As marked in the contingency table, we expect to find children in all four cells. Particularly, there would be children who pass scrambling test but fail Acc test. This will be a strong indication that the acquisition of overt case morphology is irrelevant to scrambling.

On the other hand, if the acquisition of Acc case marker is required for the acquisition of scrambling, as in the research hypothesis, we expect two possible situations; either (19a) or (19b), depending on whether Acc case acquisition is a sufficient condition or just one of several pre-requisites for scrambling.

(19) Distribution of responses under Research Hypothesis

a.	<u>Scrambling</u>			b.	<u>Scrambling</u>		
		Pass	Fail			Pass	Fail
<u>ACC</u>	Pass	√	*	<u>ACC</u>	Pass	√	√
	Fail	*	√		Fail	*	√

First, if the knowledge of Acc case marker alone is a sufficient condition for scrambling, we expect the distribution of response in (19a), where all children who pass Acc test, also pass scrambling test. This result would suggest that Acc case marker alone is a sufficient requirement to allow scrambling. However, logically, it is also possible that Acc case marker is just one of requirements, if the additional requirement has been acquired earlier than Acc case. Another possible distribution is (19b), in which some children may pass Acc test, but fail scrambling test. This would be the case in which the knowledge of Acc case marker is just one of pre-requisites for scrambling, but the additional requirement has not been acquired yet. In this case, we can be sure that there is some additional property for scrambling, and some children have not learned that property yet, leading them to fail the scrambling test.

4.4.5. Results

Originally, as many as about 50 children participated in the experiment, but only half of them yielded consistent data. This fact mainly has to do with the age factor. Our criterion for subject age was set between 2; 2 to 3; 2. Nearly all of them could pass the training test, which involves differentiating an animal doing one action from another animal doing a different action. However, when the children were presented with the actual experimental test items, which involve differentiating a pair of animals doing one action from another pair of animals doing the same action in a reversed way, many of them started to show random responses. Even though

both the training test and the actual test use the same task, it seems that many of the children experienced considerable amount of processing difficulty.

Several types of errors were observed from those children. Some children showed directional bias, by pointing out only the left side of pictures or only the right side of pictures. Some children did not keep enough attention, and ended up pointing to any picture randomly. To be on the safe side, when a subject fails the canonical word order test more than one time, we excluded the subject from analysis. 25 subjects were included for analysis. The data from them were classified as Pass or Fail for each test. The criterion for passing is getting 4 correct answers out of 5 chances, and otherwise the children were classified as fail. With this criterion the following distribution emerged.

(18) Results

		<u>Scrambling</u>	
		Pass	Fail
<u>Acc</u>	Pass	15	1
	Fail	0	9

All 25 children passed Nom test. The results above show interaction between Acc and Scrambling test. As shown in the table 9, out of 25 children 15 children passed both tests and 9 children failed both tests (See appendix for individual responses). There was one child who passed Acc test, but failed scrambling test, getting 2 answers wrong in the scrambling test. The distributional pattern in (18) shows a strong contingency between the acquisition of scrambling and the acquisition of Acc case marker (Fisher Exact test $p < .001$).

4.4.6. Discussion

The fact that all children passed Nom is not surprising given that Nom is acquired very early according to longitudinal studies. The results show a significant relation with Acc case, suggesting that overt morphological case is a syntactic condition for allowing scrambling. Now, let us discuss what these results tell us with respect to syntactic theories of scrambling. We considered three theories of scrambling: the EPP/Focus feature theory by Miyagawa (2003), the Scrambling feature theory by Gewendorf and Sabel (1998), and the Case-related theory by Bošković (2004, in press). The results argue against the scrambling feature theory since the theory does not predict any relation between Acc case and scrambling. Two syntactic theories relate overt case marker to the availability of scrambling; the EPP/Focus approach and the Case-related approach. Do the results support both theories?

In order to find it out, we need to consider the status of stimulus in the two theories. The test stimuli sentences involve scrambling in which the object is fronted to the beginning of the sentence. This type of scrambling is ambiguous between A-scrambling and A-bar scrambling both theories. Under the EPP/Focus approach, overt case morphology is a requirement for A-scrambling, but not for A-bar scrambling. This means that we expect the contingency between Acc case and scrambling only when the child uses A-scrambling to interpret the stimuli sentence. Unlike the EPP/Focus approach, the Case-related approach takes it that overt Acc case is a requirement for both A and A-bar scrambling. This means that the child needs to have knowledge about Acc case in order to interpret the stimuli sentence, regardless of which type of scrambling is employed.

Then, the question is which type of scrambling the children in the experiment were likely to use to comprehend the test stimuli sentence? Recall that Sugisaki (2000) shows that children acquire A-bar scrambling before A-scrambling. If we combine this fact with the results in my study, we expect that children are more likely to use A-bar scrambling mechanism in order to comprehend the test sentence. If that is the case, the child who has not acquired the knowledge of

overt Acc case is very likely to interpret the scrambled string as a focus adjunction structure under the EPP/Focus approach, because she does not have knowledge of overt case morphology. For this reason, the child may pass the scrambling test, but would fail the Acc test. In this case, there would be no relation between passing Acc and passing scrambling. This means that the EPP/Focus approach is in the same ball park with the Scrambling feature approach in that both theories predict that a given scrambling sentence will be acquired independently of overt case morphology, predicting the distribution of (18). This leaves us only the Case-related approach, as a theory which can account for the observed association between Acc case morphology and scrambling.

There are two remaining issues concerning the response pattern. As we have discussed in the previous section, the response pattern shown in our results is compatible with both assumptions: i) that overt Acc is the sufficient condition scrambling and ii) that an additional property is further required. Given that not all languages with an overt Acc case allow scrambling, as in Icelandic, the second interpretation is very likely. If the Case-related theory is right in assuming that the additional property is the knowledge about lack of articles, how is it plausible that the child may have learned that their language lacks articles early enough before case morphology?¹³ It is hard to find some evidence for children's knowledge of the absence of determiner category in their grammar, particularly, from spontaneous speech. However, it seems quite plausible that the child initially comes with a prior knowledge that unless they encounter the positive evidence for articles, they will not consider DPs. This predicts that the learning of the presence of articles should be easy given that any sentences they hear are likely to contain some articles. This seems to be confirmed. Ahmad Abu-Akel and Alison L. Bailey (2000), examining 17 English learning children's spontaneous speech data, reported that children, as early as 18 month, almost never make grammatical errors in the use of determiners.

¹³ Recall that for Bošković, articles are always determiners while *that*, for example, can be an adjective in some languages.

A final question we would like to consider is the difference between Turkish and Korean/Japanese. Turkish children acquire scrambling very early on, while Korean/Japanese children take quite some time. Why do we observe this difference if Acc case is the cue of scrambling? Our results offer a plausible answer to this question. As we have discussed, Turkish Acc case marks specificity. The presence of Acc is related with the semantic factor. Korean is different. Acc is totally optional as long as it is in the canonical position. This suggests that Turkish children receive a more reliable input of Acc case than Korean/Japanese children, and this facilitate their learning of Acc and consequently, scrambling.

4.5. Some Speculations

We have seen that the bulk of data argues for the Case-related theory of scrambling. Our experimental data and data from adult grammars within scrambling languages and cross-linguistic facts all point to the Case-related theory of scrambling. Now, let us pull together all the acquisitional data we have looked at. We discussed three acquisitional issues concerning the acquisition of scrambling: i) the acquisition order of A-scrambling vs A-bar scrambling, ii) acquisitional implication for A-scrambling vs MNC, and iii) positive evidence required for scrambling.

The early acquisition of A-bar scrambling with respect to A-scrambling is consistent with both the EPP/Focus theory and the Case-related theory. The contingency between Acc and Scrambling is only consistent with the case-related theory. These two acquisitional facts argue in favor of the case-related theory. However, regarding the acquisitional data about the relation between MNC and A-scrambling, there still seems to be some conflict. We saw that in Sugisaki's (2003) study, Japanese children did not use A-bar scrambling strategy to interpret the test sentences of scrambling, contrary to Sugisaki (2000). This fact is not consistent with the Case-related theory, because under the case-related theory, Japanese has both A-scrambling and A-bar

scrambling, hence the child should have been able to use A-bar scrambling strategy to interpret the test sentences. Is there any way to account for the observed acquisitional results? One way is to assume that the test items used in Sugisaki (2002) somehow require the A-scrambling strategy. In other words, one could speculate that instances of multiple scrambling that Sugisaki tested are possible with A-scrambling. If this is on the right track, one has to say that apparent multiple scrambling in only A-bar scrambling is not actually multiple scrambling, but single phrase scrambling plus topicalization/focalization.

4.6. Implications for the Theory of Parameter Setting

The experimental results also have some implications for the theory of parameters.

Chomsky (1981:6) conceives of parameters as possible options as to how some principles of grammar apply. The interactions of certain possibilities of parametric variation are expected to yield collections of complex properties. Accordingly, setting of a single parameter may have wide reaching effects on various parts of the grammar, which may look superficially unrelated.

For instance, Chomsky (1981) used the ‘pro-drop parameter’ for an example of the proliferating effects of a single parameter setting. He assumes that the pro-drop parameter is responsible for collections of grammatical properties: null subject, lack of that-trace effect, subject-inversion, etc. Kuroda (1988) assumes that languages have options as to whether INFL undergoes agreement or not. He claims that the choice of this parametric value has effects on a number of other syntactic properties of the grammar such as the availability of ‘multiple scrambling’ and ‘multiple case constructions’.¹⁴ Language variation with respect to scrambling was also analyzed in the similar vein. Hale (1980) proposes that languages are parameterized as to whether the government principle ‘clicks’ in a given language or not. In languages where the government principle clicks, subject and object have to occupy unique positions, hence requiring

¹⁴ Note that the bi-directional implication between multiple scrambling and multiple case constructions cannot be correct. All scrambling languages allow multiple scrambling, but not necessarily multiple case constructions.

strict positions of arguments. However, languages in which the principle is not operative do not require the strict positioning. As a result, generation of arguments is more flexible, allowing the scrambling phenomenon. This original formulation of parameters seems to highlight the broad consequences of a single parameter setting, capturing the clustered properties that are found together in a given language. It has been taken as a promising line of language acquisition research, which can account for the relative rapidity and the efficient success of children's language development.

With the introduction of the Minimalist Program (Chomsky 1995), Generative Grammar has undertaken a considerable change to the syntactic theory, reducing a vast number of assumptions adopted in the previous model, Government Binding framework. One of the main ideas of Minimalist Program concerns the notion of syntactic operation. Unlike the previous mechanism 'Move α ', which allows movement anywhere anytime, Minimalist Program theory assumes that syntactic operations must take place for a reason, particularly, checking morphological properties in inflectional categories. Given this, syntactic variation is considered to be a result of variations in morphological properties of non-substantive parts of the lexicon. This approach to parameterization is often called 'lexicalist' hypothesis of parameters (Borer and Wexler 1994). Chomsky (1995: 198) exploits this idea to account for language variation in V-Raising. Assuming that the location of the parameter is a V-feature in a functional head Agr, he argues that English has Agr with [weak V-feature], allowing LF checking, while French has Agr with [strong V-feature], forcing overt V-Raising. The difference in the feature strength is assumed to be related to richness of verbal morphology.

The present study provides a supporting argument for the lexicalist hypothesis of parameters. Our results have shown that overt Acc case morphology affects the word order domain of the grammar; the presence of overt case morphology makes scrambling available. It confirms that a difference in a function category affects a seemingly unrelated property of the grammar, keeping the original spirit of parameters in Chomsky (1981).

According to the Case-related theory by Bošković (2004, in press), the overt case marker plays a function in satisfying formal licensing of an argument (the principle of Argument Identification Requirement), which allows merge in a non-theta position, giving rise to scrambling effect on the surface. This means that the cross-linguistic variation with respect to the availability of scrambling is derivable from interactions with the relevant principles in UG. In this approach, the availability of scrambling is determined by overt case morphology, lack of the functional category DP and syntactic consequences that follow from the relevant UG principles. This analysis removes the scrambling parameter from the theory of grammar, which seems to be a desirable move.

4.7. Implications for the Theory of Learnability

Our result that the child acquires scrambling using case morphology makes an important implication for the learnability theory, particularly concerning how the learner maps linguistic experience to a particular grammar. For the last couple of years, language acquisition research has seen the beginning of more refined proposals of input-oriented learning, under the name of ‘*statistical learning*’ model (Saffran, Aslin & Newport 1996, Aslin, Saffran & Newport 1998, Newport et al 2000). This approach assumes that the learners have a highly complex data processing mechanism which allows them to compute statistical property of input. The main domain of application of this approach so far is the area of learning of word boundary. They argue that even young infants (as early as 8 months) can in fact perform remarkably complex computations of co-occurrences among adjacent elements. One statistical property that they have investigated with various experimental studies is ‘transitional probabilities’, which refers to the probabilities of co-occurrence of adjacent syllables within words and adjacent syllables across words. They claim that infants use the transitional properties to figure out word boundaries, and further argue that such ability is not specific to language. With this achievement, they are trying

‘Does she always read the newspaper?’

(19a/b) are the cues for verb-raising because the structural representations of (19a/b) unambiguously require verb movement to T. When children encounter such a cue reliably, this will trigger the acquisition of verb-raising. In a similar vein, a cue for verb second grammar is a structure commencing with a XP followed immediately by a finite V, where there is no fixed grammatical or thematic relation between the initial phrasal category and the finite verb. This requires analyzing the XP as in Spec CP, so that the XP in spec CP is the cue for the positive value of the verb second parameter. As for scrambling, given that the positive evidence for a syntactic cue available to the child was extremely poor, we can assume that the only reliable cue which seems to be accessible to the child is a morphological cue; the overt case morphology,

Lightfoot argues that the cue for a particular parameter is specified in UG along with each parameter. In other words, he assumes that UG specifies not only a set of parameters, but also for each parameter a cue, or multiple cues. However, as for scrambling, if we adopt Boskovic’s scrambling theory, there is no such need of specifying the cue in UG, since there is not scrambling parameter any more, and further the availability of scrambling is derivable from the interaction between UG principles and functional morphology.

4.8. Conclusion

For the last few decades, the UG-based approach within the Principles and Parameters framework (Chomsky 1981) has undergone several considerable transformations, building more detailed and highly elegant model of the language faculty (Chomsky 2000, 2002, 2003). In contrast, in the language acquisition research, the strongest argument for the innateness hypothesis based on the disparity between input and output has begun to be questioned, or at least re-evaluated by alternative approaches which put more emphasis on the role of input. The

findings of invariant properties of languages are no longer considered to be self-evident for the innateness hypothesis as the source of that knowledge.

In light of this current trend, this study makes a significant contribution to the theory of language acquisition. We have seen that children's production data almost mirror the adult speech pattern, seemingly providing support for input-driven learning approach; scrambling utterances were almost zero in both the child speech and the adult speech. However, the investigation of the acquisition of scrambling by an experimental study reveals to us that the child learns scrambling using the seemingly unrelated property, overt case morphology. This provides strong evidence against input-driven learning theory since there is no input for scrambling sentences that the child could learn from. Our results show that the child must have innate knowledge about the abstract link between scrambling and Acc case, providing supporting evidence for the UG-based learning approach.

Appendix to Chapter 4

A. Individual Responses

	Name	age	A1	B1	C1	D1	B2	C2	D2	A2
1	JB	3;2	C	C	C	C	C	C	C	C
2	HI	2;2	C	C	C	C	C	C	W	C
3	YB	2;7	C	C	W	C	C	W	W	C
4	JY	2;5	C	C	C	W	C	C	C	C
5	SY	2;9	C	C	C	C	C	C	C	C
6	HA	2;5	C	C	W	W	C	C	C	C
7	DH	3;0	C	C	W	C	W	C	C	C
8	CH	2;8	C	C	W	W	C	C	W	C
9	SU	2;6	C	C	W	W	C	C	W	C
10	SB	2;5	C	C	C	W	C	W	W	C
11	JK	3;1	C	W	C	C	C	C	C	C
12	YS	3;2	C	C	C	C	C	C	C	C
13	CY	3;0	C	C	C	C	C	C	C	C
14	HJ	2;6	C	C	C	C	C	C	W	C
15	JI	2;9	C	C	W	W	C	W	C	C
16	JB	3;0	C	C	W	C	C	W	W	C
17	YO	2;7	C	C	W	C	C	W	W	C
18	MS	3;1	C	C	W	W	C	C	W	C
19	LS	2;10	C	C	W	C	C	C	W	C
20	JY	3;2	C	C	C	C	W	C	C	C
21	YK	3;0	C	C	C	C	C	C	W	C
22	SM	2;10	C	C	C	C	C	C	W	W
23	DA	3;1	C	C	C	C	C	C	C	C
24	JN	3;0	C	C	C	C	C	C	W	C
25	CL	3;3	C	C	C	C	C	C	C	C
	Mean	2;10								

	Name	age	B3	C3	A3	D3	B4	D4	A4	C4	D5	A5	C5	B5
1	JB	3;2	C	C	C	C	C	C	C	C	C	C	C	C
2	HI	2;2	C	C	C	C	C	C	C	C	C	C	C	C
3	YB	2;7	C	W	C	W	C	W	C	C	C	C	W	C
4	JY	2;5	C	C	C	C	C	C	C	C	C	C	C	C
5	SY	2;9	C	C	C	C	C	C	C	C	W	C	C	C
6	HA	2;5	C	C	C	C	C	C	C	C	C	C	C	C
7	DH	3;0	C	W	C	W	C	W	C	C	W	C	W	C
8	CH	2;8	C	W	C	W	C	W	C	C	W	C	W	C
9	SU	2;6	C	W	C	W	W	W	C	W	C	C	C	C
10	SB	2;5	W	C	C	C	C	C	C	C	C	C	C	C
11	JK	3;1	C	C	C	C	C	C	C	C	C	C	C	C
12	YS	3;2	C	C	C	C	C	C	C	W	C	C	C	C
13	CY	3;0	C	C	C	C	C	C	C	C	C	C	C	C
14	HJ	2;6	C	C	C	C	C	C	W	C	C	C	C	C
15	JI	2;9	C	W	C	W	C	W	C	W	W	C	C	C
16	JB	3;0	C	W	C	W	C	C	C	C	W	C	W	C
17	YO	2;7	C	W	C	W	C	W	C	C	C	C	W	C
18	MS	3;1	C	W	C	C	C	C	C	W	W	C	C	C
19	LS	2;10	C	C	C	C	C	W	C	W	W	C	W	C
20	JY	3;2	C	C	C	C	C	C	C	C	C	C	C	C
21	YK	3;0	C	C	C	C	C	C	C	C	C	C	C	C
22	SM	2;10	C	C	C	C	C	C	C	C	C	C	C	C
23	DA	3;1	C	C	C	C	C	C	C	C	C	C	C	C
24	JN	3;0	C	C	C	C	C	W	C	C	C	C	C	C
25	CL	3;3	C	C	C	C	C	C	C	C	C	W	C	C
	Mean	2;10												

- A: Canonical order test
- B: Nom test
- C: Acc test
- D: Scrambling test

B. Test items

- A1. dongmulwon-e horangi-ga issesseyo, kurende ku horangi-ga wonsungi-lul mulesseyo
 zoo-in tiger-Nom was then that tiger-Nom monkey-Acc bit
 'There was a tiger in the zoo, then that tiger bit a monkey.'
- B1. yeonmos-e kepwuki-ga issesseyo, kurende ori-ga ccosagasseyo.
 pond-in turtle-Nom was then duck-Nom chased
 'There was a turtle in the pond. Then a duck chased the turtle.'
- C1. kongwon-e tokki-ga issesseyo, kurende daramci-lul milesseyo.
 park-in rabbit-Nom was then squirrel-Acc pushed
 'There was a rabbit in the park, then the rabbit pushed a squirrel.'
- D1. supsok-e wonsungi-ga issesseyo, kurende ku wonsungi-lul kom-i mulesseyo.
 wood-in monkey-Nom was then that monkey-Acc bear-Nom bit
 'There was a monkey in the wood, then a bear bit the monkey.'
- A2. donmulwon-e saza-ga issesseyo, kurende ku saza-ga kepwuki-lul ccosagasseyo.
 zoo-in lion-Nom was then that lion-Nom turtle-Acc chased
 'There was a lion in the zoo, then the lion chased a turtle.'
- B2. yeonmos-e tyoe-ci-ga issesseyo, kurende songaci-ga milesseyo.
 pond-in pig-Nom was then cow-Nom pushed
 'There was a pig in the pond, then a cow pushed the pig.'
- C2. supsok-e daramci-ga issesseyo, kurende tokki-lul mulesseyo.
 wood-in squirrel-Nom was then rabbit-Acc bit
 'There was a squirrel in the wood, then the squirrel bit a rabbit.'
- D2. kongwon-e kyrin-i issesseyo, kurende ku kyrin-ul wonsungi-ga ccosagasseyo.
 park-in giraffe-Nom was then that giraffe-Acc monkey-Nom chased
 'There was a giraffe in the park, then the a monkey chased the giraffe.'
- A3. donmulwon-e tokki-ga issesseyo, kurende ku tokki-ga kangaci-lul milesseyo.
 zoo-in rabbit-Nom was then that rabbit-Nom dog-Acc bit
 'There was a rabbit in the zoo, then the rabbit bit a dog.'
- B3. supsok-e kom-i issesseyo, kurende sasum-i ccosagasseyo.
 wood-in bear-Nom was then deer-Nom chased
 'There was a bear in the wood, then a deer chased the bear.'
- C3. kongwon-e tyci-ga issesseyo, kurende koyangi-lul milesseyo
 park-in pig-Nom was then cat-Acc bit
 'There was a pig in the park, then the pig bit chased a cat'
- D3. supsok-e songaci-ga issesseyo, kurende ku songaci-lul sasum-i milesseyo
 wood-in cow-Nom was then that cow-Acc deer-Nom pushed
 'There was a cow in the park, then a deer pushed the cow'
- A4. kongwon-e kangaci-ga issesseyo, kurende ku kangaci-ga tokki-lul milesseyo
 park-in dog-Nom was then that dog-Nom rabbit-Acc pushed
 'There was a dog in the park, then the dog pushed a rabbit'
- B4. donmulwon-e horangi-ga issesseyo, kurende kepwuki-ga mulesseyo.

- zoo-in tiger-Nom was then turtle-Nom bit
 'There was a tiger in the zoo, then a turtle bit the tiger.'
- D4. supsok-e wongsung-ga issesseyo, kurende tyeci-lul ccosagasseyo.
 wood-in monkey-Nom was then pig-Acc chased
 'There was a monkey in the wood, then the monkey chased a pig.'
- C4. yeonmos-e mal-i issesseyo, kurende ku mal-lul kepwuki-ga milesseyo
 pond-in horse-Nom was then that horse-Acc turtle-Nom pushed
 'There was a horse in the pond, then a turtle pushed the horse.'
- A5. supsok-e daramci-ga issesseyo, kurende daramci-ga kangaci-lul ccosagasseyo.
 wood-in squirrel-Nom was then that squirrel-Nom dog-Acc chased
 'There was a squirrel in the wood, then the squirrel chased a dog.'
- B5. donmulwon-e tokki-ga issesseyo, kurende kepwuk-ga milesseyo.
 zoo-in rabbit-Nom was then turtle-Nom pushed
 'There was a rabbit in the zoo, then a turtle pushed the rabbit.'
- C5. yeonmos-e koyangi-ga issesseyo, kurende tokki-lul milesseyo
 pond-in cat-Nom was then rabbit-Acc pushed
 'There was a cat in the pond, then the cat pushed a rabbit.'
- C5. donmulwon-e saza-ga issesseyo, kurende ku saza-lul kirin-i ccosagasseyo.
 zoo-in lion-Nom was then that lion-Acc giraff-Nom chased
 'There was a lion in the zoo, then a giraffe chased the lion.'

Chapter 5. Concluding Remarks

5.0. Summary

This thesis examined acquisition of language particular syntactic properties in Korean to explore the nature of the mechanisms used in the acquisition process. The investigation was couched in the discussion of two opposing views to language acquisition: the input-driven learning approach and the UG-based learning approach.

In the introductory chapter, we discussed several previous acquisitional findings with respect to the two different learning approaches, and reviewed major motivations of the UG-based learning approach. Various phenomena of mismatch between the hypotheses that the child entertains and their input patterns led us to postulate a mediator to bridge the acquisition process and the input patterns, supporting the UG-based learning approach.

In Chapter 2, 3 and 4, we investigated two areas of grammar which are subject to cross-linguistic syntactic variation: multiple Nominative Case marking (MNC) and scrambling. In contrast to the implicit assumption in the acquisition literature that the language input available to the child is fully representative of samples of all the relevant properties possible in the adult grammar, we found a discrepancy between the adult grammar and the actual input available to the child. Korean input data presented a poverty of stimulus situation. The positive evidence for both MNC and scrambling was drastically rare. This surprising situation invited us to explore how children figure out the availability of the two grammatical properties in such impoverished input.

Chapter 2 investigated the acquisition of MNC. We started with a simple acquisition scenario that the child may learn MNC if they know Nom case as a marker for subject and the peculiarity of stative predicates requiring Nom case for the internal argument. The observation from longitudinal production data, however, revealed that children do not produce MNC of the

type with psych predicates even after they have learned the two things, suggesting that the lack of MNC in the child speech may be due to the learning of some additional grammatical property of MNC.

We considered two syntactic theories of MNC, which treat multiplicity of Nom case as a grammatical property for MNC. The Default case marking approach (Fukui and Takano 1998) assumes MNC as a result of Korean Nom case being a generalized default case, while the Multiple feature checking approach (Ura 1996, 1999, 2000) takes that multiple Nom is a parameterized property of the functional category T as to the possibility of multiple feature checking. Along with the two syntactic theories of MNC, we also considered the possibility that somehow the lack of MNC may be explained by the poor input frequency of MNC. The child may have relevant knowledge of MNC, but simply may not have produced it, assimilating to their input patterns.

In order to tease apart these possible interpretations as well as to evaluate the two different syntactic theories of MNC, we carried out an experiment. In the experiment, we used MNC of inalienable relation, which did appear in the input about 10% (out of all stative predicates), and compared them with Dative subject (DS) sentences, which were non-existent in the input. This comparison allowed us to test three hypotheses. If input frequency determines the order of acquisition, MNC should be acquired earlier than dative subject sentences since there were 10% of MNC, but no instances of Dative subject sentences. The Default case approach predicts that children will acquire not necessarily later than DS sentences. On the other hand, multiple feature checking approach predicts that the child will acquire MNC after Dative subject sentences because they need to set the multiple feature checking parameter to the positive value.

The results showed that all children passed DS test, while some children failed MNC test. This result clearly argues against the input-driven learning approach since children acquired DS sentences earlier even if the frequency of MNC sentences was much higher than DS. This empirical data supports Ura's (1996, 1999, 2000) multiple feature checking approach over the

Default case approach because the fact that children acquire MNC later than DS can be explained by the time required for the setting of the multiple feature checking parameter.

Chapter 3 and 4 explored acquisition of scrambling. In chapter 3, we reviewed syntactic theories of scrambling: the EPP/Focus approach (Miyagawa 1999), the Scrambling feature approach (Grewendorf and Sabel 1999), and the Case-related approach (Bošković 2003). We discussed two acquisitional implications of each scrambling theory: one issue concerning the acquisition order between A-scrambling and A-bar scrambling, and the other issue concerning the acquisitional relation between A-scrambling and MNC. We saw that the three scrambling theories make some overlapping predictions for the two acquisitional issues.

In chapter 4, we turn to a learnability question of whether each theory of scrambling guarantees the positive evidence for the learning of scrambling. From the study of input data, we discovered that scrambling utterances are extremely rare, suggesting that children are not likely to learn scrambling by hearing scrambling utterances. In order to solve the learnability problem, we proposed that the child may use some kind of trigger, which may indirectly indicate the availability of scrambling. Based on the case dropping constraint and cross-linguistic facts, we hypothesized that Acc case may be a possible cue for scrambling.

We conducted an experiment to test the acquisitional association between scrambling and Acc case marker. The results showed that there is a strong contingency between passing Acc and passing scrambling, suggesting that the child uses Acc case to acquire scrambling. This result, together with previous acquisitional facts concerning the acquisition order between A-scrambling and A-bar scrambling, supports the Case-related scrambling theory by Bošković (2003).

The empirical data from both sets of studies showed that the Principles and Parameters learning model of UG-based approach is indispensable to account for the acquisition of language particular properties. A parameter theory of Multiple Nom can explain the time required for the acquisition of multiple Nom. The prior knowledge of abstract link between scrambling and Acc can explain how children acquire scrambling on the near absence of scrambling sentences.

5.1. Significance of Complementing Longitudinal Production Data with Experimental Studies

In this study, we demonstrated a significance of complementing longitudinal production data with experimental studies in studying the nature of language acquisition. We observed from longitudinal production data of Korean that children rarely produced MNC and scrambling utterances. When we examined their parental input, we found the similar pattern. Input frequencies of both MNC and scrambling were very low. The close resemblance between the child speech and the adult speech in the longitudinal production data gave us the appearance that language acquisition proceeds by input-driven learning model. Indeed, often times in the language acquisition literature, this kind of frequency match between the child speech and the adult speech in the production data is used as supporting evidence for input-driven learning approach (Chung 1999). We pointed out that such an input-matching cannot account for learnability problem. It cannot explain how children eventually come to acquire scrambling in spite of the near absence of scrambling utterances in the input.

We further noted that the absence of some linguistic expression in the production data may be caused by different reasons. The child actually may not have the relevant grammatical knowledge. Alternatively, the child may have the knowledge, but simply chose to not to produce the relevant constructions, because they were tuning to the adult speech patterns. Our experimental results of MNC showed that the absence of MNC in the child speech was indeed due to the lack of grammatical knowledge concerning multiplicity of Nom case, rather than just the matching of input frequency. As for scrambling, the experimental results showed that the absence of scrambling in the child speech was due to assimilating to the adult speech patterns, rather than due to the lack of grammatical knowledge of scrambling because some of the children studied in the production data were producing Acc case at adult-like levels. These facts

demonstrate that the experimental study can provide important data to tease apart the different possibilities as well as help to explore learnability problem.

5.2. Optionality vs Obligatoriness in Language Acquisition

We provided an argument for postulating language specific innate knowledge of language from the comparative study of developmental contours. In investigating the developmental contours of Nom and Acc case, we found a striking contrast between the two cases Nom and Acc in their developmental pattern; Nom case showed a relatively rapid change with steep increase to adult level of frequency, while Acc showed a gradual change. Further, children were able to match the adult frequency of Acc case, which was only 20 to 30%.

At first sight, the remarkable input matching of Acc case seemed to provide strong evidence for input-matching learning model. We compared this case with another case of input inconsistency studied by Newport and Aslin (2000), in which deaf children, who were exposed to inconsistent morphemes of ASL verbs of motion from their non-native hearing parents, regularized the use of the morphemes up to 90% of the times. Newport and Aslin hypothesized that children have a tendency to regularize a major pattern they find in the input, assuming that that is how children use statistical information to form rules of grammar. We argued that Newport and Aslin's hypothesis cannot be correct, nor the input matching model account. Newport and Aslin's hypothesis incorrectly predicts that Korean children will regularize the Acc case up to 90%, as in deaf children's case. The input-matching model account also makes a wrong prediction for deaf children's case because it incorrectly predicts that the deaf children would use the relevant morphemes only with the adult level of frequency.

We proposed that the only way to account for the apparently contrasting results for the similar case is to assume that children have a prior knowledge about what is obligatory and what can be optional in the grammar. We argued that children know that verb agreement is obligatory,

and this prior knowledge enables him to produce the morphemes in all obligatory contexts, regardless of the inconsistent input.

A similar argument was applied for the difference in the development between Nom and Acc. We argued that the child comes with a prior knowledge about which case is obligatory, and which case is optional. If a case is obligatory, the child can produce it in obligatory contexts quickly after they learn the case form. If a case is optional, the child will have to pay attention to the adult frequency, and this brings them to slow down in the development, resulting in the gradual change, and also allowing them to stay in the input-rate.

Our argument that the innate knowledge specifies agreement to be obligatory makes some predictions regarding children's acquisition of verbal agreement. It predicts that children learning languages with verbal agreement will show the similar behavior as the case of the deaf children's learning of morphemes of ASL verbs of motion. The prediction seems to be borne out in subgroup of languages, which are characterized as pro-drop languages. Salustri and Hyams (2003) report that the study of acquisition of Catalan and Spanish show that around 98% of verbs are inflected with agreement almost from the start.¹

However, there seems to be some counter-evidence from a class of languages characterized as non-pro-drop languages. It is reported that children learning those non-pro-drop languages go through a stage of optional verb agreement, in which they alternate uninflected verbs with inflected verbs in main clauses, a phenomenon known as 'Optional Infinitives' (OI) (see Wexler 1994, Platzack 1992, Haegeman 1995, Hamann and Plunkett 1998). This raises a question; if the child comes with a prior knowledge that agreement cannot be optional, as we have argued, why would they produce the agreement only optionally? There are three main analyses of OI: Wexler's (1999) tense omission model, Rizzi's (1993, 1994) truncation

¹ Salustri and Hyams report data from Bel (2001), which shows that 3 Catalan and 3 Spanish-speaking children with age range from 1;7-3;0 produced uninflected verbs only less than 4%.

hypothesis and aspect hypothesis by Salustri and Hyams's (2004).² Here we contrast two views, which seem maximally different from each other: the tense omission model and the aspect hypothesis.

Wexler (1999) proposes that infinitive verbs originate due to underspecification of tense feature. He argues that the OI stage of the child grammar has a so-called 'Unique Checking Constraint', which allows that the D-feature of a DP can check the D-feature of either T or Agr, but not both.³ In adult grammar, there is no such constraint; a subject enters into a specified-head relation with both T and Agr.⁴ However, in child grammar a representation in which the subject checks the D-feature of both T and Agr violates the Uniqueness Constraint. To avoid this violation, children leave T underspecified, violating another constraint, 'Tense constraint', which requires that a main clause must include a specification of tense. Consequently, TP is removed from the clausal representation. How does this theory explain the alternation with finite verbs? Wexler argues that the child can also violate the uniqueness constraint, and preserve tense constraint instead, because in their grammar, either one violate the same number of constraint.⁵ Wexler's account for OI seems to contradict our argument. If OI is indeed a result of not projecting tense feature because the constraint governing checking of agreement feature is

² Rizzi's (1996) hypothesis is based on an assumption called an 'axiom on clausal representation' that states that CP is the root of all clauses. Rizzi proposes that the axiom is always operative in adult grammar, but applies optionally in child grammar. When the axiom is not operative, the child grammar can generate structures truncated below CP. In this theory, OI clauses are the structures truncated below TP; that is, they are VPs. Since they are VPs, they cannot encode the temporal information, generating non-finite verbs. In Wexler's (1999) tense omission model, uniqueness constraint is responsible for the underspecification of Tense feature, while in Rizzi's truncation model the deficiency of tense is tied to a different hypothesis on which node can count as the root of the clauses. Both Wexler's tense omission model and Rizzi's truncation hypothesis assumes that tense is somehow deficient in children's grammar. In this sense, both theories seem to be incompatible with my argument because they eventually allow optional projection of tense.

³ 'Uniqueness constraint' states that 'A subject can check the uninterpretable feature of either T or Agr, but not both.' (Wexler 1999).

⁴ Wexler (1999) assumes that both Agr and T have a D-feature, which must be eliminated by being checked against the D-feature of a DP subject that raises to Spec TP and Spec AgrP.

⁵ This possibility is introduced as 'Minimize Violations', stated as below.

Minimize Violations

'Given two representations, choose that violates as few grammatical constraints as possible. If two representations violate the same number of constraints, then either one may be chosen' (Wexler 1999).

violable, the creolization effect with ASL learning deaf children seems to be puzzling. What motivated them to produce the agreement morphemes in nearly all obligatory contexts, in spite the fact that they received the input of only 20-40% correct use? Wexler's analysis seems to leave this unexplained.

Salustri and Hyams' (2003) aspect hypothesis seems to avoid the problem. They argue that OIs are not instances of leaving out inflectional agreement optionally. This shows that the distribution of OI has correspondence with some semantic interpretation. They observe that infinitives are mostly used with eventive verbs, and have modal/future meaning (Clashen 1986, Becker and Hyams 1999, Lasser 2000).⁶ This means that their use of infinitive forms is not the case of omitting agreement morphemes; hence, there is no issue of optionality of agreement.

5.3. Statistical Learning Used in Frequency Matching

In this study, we reported the acquisitional data which showed that children succeeded in matching the adult level of frequency. In particular, children's matching 20-30% rate of Acc use was quite remarkable. How does the child zero in on the exact rate? Does the child have the access to overall frequency of Acc case in the input? We suggested that they do, as argued by statistical learning model (Saffran et al 1996, Newport and Aslin 2002). Here we provide more evidence that seems to point to the same direction.

It is known that the rate of the use of null-subjects varies among pro-drop language. Table 1 shows percentage of the rate of overt subjects in several pro-drop languages. Kim (2000) observes that regardless of their target languages, children acquiring any language initially produce overt subject less often than adults of their language community.

⁶ Salustri and Hyams (2004) cite the data of German children reported by Berger-Morales, Salustri and Jikerson (2003); 89% of the OIs were of future and modal interpretation, and only 10% of OIs were of present and past tense interpretation.

(1) Table 1. Cross-linguistic variation in the rate of overt subjects

(Kim 2000)

Language	The rate of overt subject in input
Korean	35-45%
Japanese	26-38%
Mandarin	64%
Cantonese	58%
Italian	46-56%
Brazilian Portuguese	56%

She reports that nevertheless, at a fairly young age, children not only know whether or not their target language allows null subjects, but also show sensitivity to actual frequencies of overt subjects in their target language. Table 2 shows the age when they reach adult-rate of the use of overt subject. The age with Mandarin and Chinese is very late compared with other languages. This seems to be related to a different method in data collection.⁷ Apart from these two languages, it seems that before the age of 3, they match the adult input. Now, the question is how they manage to match it.

⁷ According to Diane Lillo-Martin (p.c.), the data was collected in story telling context, and may not be compatible with other longitudinal speech data even though it will be consistent within-data analysis.

(2) Table 2. The age of adult-rate use of overt subjects

(Kim 2000)

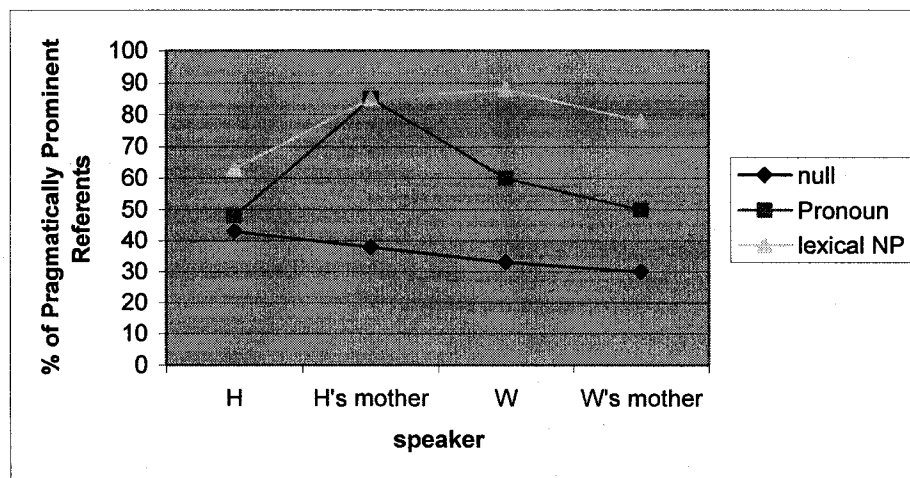
Language	The age of adult-rate use of overt subject
Korean	2;1-2;4
Japanese	1;6-2;6
Mandarin	4;1-4;4
Cantonese	3;4
Italian	2;5~
Brazilian Portuguese	2;3-2;10

Clancy's (1993) study provides more detailed information on this phenomenon. She studies the patterns of argument realization in child speech of Korean. In Korean, as in other pro-drop languages, an argument can be realized in 3 different types; null, overt pronoun, lexical NP. She analyzes distributions of the 3 types of argument in child speech and their input, using a notion of pragmatic prominence defined as follows: i) Queries: the referent is being queried or is an answer to a question, ii) Contrastiveness: the referent is one of at least two potential referents in the context to whom the same predicates is or could be applied, iii) Lack of (immediate) prior mention: the referent is being introduced into the discourse for the first time, or has been mentioned before but not in the prior clauses, iv) Absence: the referent is not in view in the room at the time of utterance.

She searches each instance of these 3 types of subjects, and checked whether each argument has one of the features of pragmatic prominence defined as above. She found a surprising fact. One child almost mirrored the adult patterns in matching the proportions of each argument used with pragmatic prominence. The results are given in the figure 1.

(3) Figure 1. Pragmatic Differentiation of Referential Options

(Clancy 1993: 312)



Age of the child H: 1;7-2;1

W: 1;10-2;3

There is some complication with child H's data. He does not seem to show a clear distinction among three types of arguments.⁸ Child H's mother does not seem to distinguish pronoun vs lexical NP in terms of the pragmatic prominence, either. On the other hand, child W almost mirrors the pattern of his mother. Most use of lexical subjects was pragmatically prominent, and about 50% of the pronouns were pragmatically prominent. The majority of null subjects were not pragmatically prominent.

The above data shows that the child seems to use the same principles that govern the relation between linguistic forms and their functions in the pragmatic use. A further surprising fact is that the child manages to obtain the similar level of proportion of each type of argument used with pragmatic prominence. This fact, together with the frequency match in the acquisition of Korean Acc case, leads us to conclude that children indeed have access to the overall frequency of each type of argument used with pragmatic prominence.

⁸ Clancy (1994) reports that the child H was not less advanced compared with the child W. The child made some errors in which sometimes he used lexical subject when the null subject is more appropriated, etc.

Let us ask how this conclusion fits in the big picture of the language acquisition. We have seen from the investigation of the acquisition of MNC and scrambling that children do not use frequency information to acquire the syntactic properties, even though production data seems to mirror the frequency patterns of their parental input. However, we have found the evidence indicating that children show remarkable performance in matching the frequency of the input, as advocated by a recent learning theory such as statistical learning (Saffran et al 1996, Newport and Aslin 2002). Is this a problem?

We argue that it is not. Consider the status of frequency in the grammar. It is clear that how often a certain linguistic expression is used in the performance has nothing to do with its grammatical status. It may be used very rarely, but perfectly grammatical. Grammar is a system of mental knowledge which determines a given sentence either to be grammatical or ungrammatical. Given this, it is not surprising that children did not rely on frequency in acquiring syntactic properties. The areas we found frequency to play a role was the areas of performance in which they somehow figured out how often a given expression is used in the input, not the areas in which they have to learn whether a given property is grammatical or not. This conclusion makes a significant point to the current trend of language acquisition research, because it embraces the argument of statistical learning model regarding the point that children are indeed capable of performing complex statistical computations, but at the same time shows the limitation of such learning model. We showed that such statistic computational ability plays a role in matching the aspects of performance in the input, not in acquiring syntactic properties.