

# EARLY ACQUISITION OF WORD ORDER: EVIDENCE FROM HINDI-URDU AND MALAYALAM

MAYA LEELA

A doctoral dissertation submitted for the degree of  
*Philosophiæ Doctor in Cognitive Science and Language*

Supervised by Dr. Anna Gavarró



Universitat Autònoma  
de Barcelona

Centre de Lingüística Teòrica  
Departament de Filologia Catalana  
Universitat Autònoma de Barcelona

Maya Leela: *Early Acquisition of Word Order: Evidence from Hindi-Urdu  
and Malayalam* © July 2016

Dedicated to Unniyamma.

## ABSTRACT

---

The acquisition of word order in two under-researched languages, Hindi-Urdu and Malayalam, is explored under the assumptions of Universal Grammar and Very Early Parameter Setting. The acquisition of the OV order was tested through an experiment that was conducted on 19-month old native Hindi-Urdu infants using a combination of the preferential-looking mode, the weird-word-order paradigm, and pseudo-verbs. The results showed that the infants parsed the grammatical SOV order and did not parse the ungrammatical VSO order, indicating that the parameter responsible for the OV order was set at 19 months of age. The acquisition of scrambling was investigated by means of three experiments in Malayalam speaking children. The first experiment tested the acquisition of the canonical word order SOV, and the scrambled non-canonical word orders OSV, OVS, and SVO, with no discourse context, through a picture matching task in children aged 3 to 5-years-old. The experimental results showed above chance performance in the comprehension of the canonical and most scrambled word orders by all age groups. The second experiment tested the acquisition of the canonical and the non-canonical word orders with a discourse context, in an act-out task in 2-year-old children. The results revealed at ceiling performance by all children. The third experiment, run with 2-year-old children, tested the acquisition of canonical and scrambled sentences with an anaphor. All children performed above chance indicating that anaphor resolution is adult like. Usage-based theories and the implications of the theory of Relativized Minimality in child grammar are critically analyzed using the results of the Hindi-Urdu and Malayalam experiments. The evidence gathered is in support of the fact that adult-like competence of abstract syntax is present within child grammar at early ages and that movement appears to be adult-like.

## PUBLICATIONS

---

Some of the results in this thesis have appeared or are to appear in the following publications:

1. Gavarró, A., Leela, M., Rizzi, L., & Franck, J. (2015). Knowledge of the OV parameter setting at 19 months: Evidence from Hindi–Urdu. *Lingua*, 154, 27–34.
2. Gavarró, A., Leela, M. Child Relativized Minimality and the acquisition of word order in Malayalam. To appear in *Linguistic Analysis*.

“A little parameter parametrizes the whole language”

— (Baker, 2002, p.44)

## ACKNOWLEDGMENTS

---

I would like to express my sincere gratitude to Dr. Anna Gavarró, my supervisor, whose selfless dedication to guide me in producing an excellent and impeccable result is the main reason for the final shape of this thesis. Anna has given me continuous support with her most updated knowledge on theoretical frameworks and experimental research, and her most sensible approach towards life. I thank all her contributions to me and my thesis.

I thank the Departament de Filologia Catalana and the Centre de Lingüística Teòrica at the UAB for accepting me as a doctoral candidate and allowing me to carry out the research. I formally express my gratitude to the Agencia Española de Cooperación Internacional para el Desarrollo and the Ministerio de Asuntos Exteriores y de Cooperación for the MAEC-AECID funding provided for the research.

The completion of the Hindi-Urdu experiment became possible with the help of the amazing ladies at the Associació Cultural, Educativa i Social Operativa de Dones Pakistaneses (ACESOP). I thank the president and the founder of the organization, Dr. Huma Jamshed, and the members who participated in the recording of the test sentences, as well as the members who helped me find the subjects. I especially thank all the mothers who managed to get permission from the patriarchy bound environments and come with their infants for the testing.

For all the technical knowledge I gathered in handling the eye-tracker, I thank both David Aguilar and Juan Manuel Garcia at the Departament de Psicologia Bàsica, Universitat de Barcelona, and Miquel Serra for giving me permission to use the lab and the equipment.

I thank the students and teachers at the Wisdom Valley Kindergarten, and Ananthapuri School at Trivandrum for their participation in the Malayalam experiments. The drawings used were produced by the artist

Justyna Kamykowska, and I thank her for allowing me to use the copyrighted materials. I thank Arathy Chandini, Vaisakh Surendran and Ahana Balan for helping me with the data collection.

I would like to thank Luigi Rizzi, Julie Franck and Akira Omaki for giving me an opportunity to collaborate with their acquisition research. I thank Anna Espinal at the Servei d'Estadística Aplicada, UAB and Andrés Posada at the Université de Genève for the statistical analysis.

Words of gratitude are not enough to thank Ivo Pletikosić, physicist from Croatia, for all the help in overcoming the hurdles I faced with the *classicthesis* template and with LyX. He replied quickly to all my emails, patiently guided me through every step and helped me immensely to produce this thesis in this form. Thanks a lot, Ivo, you being in the LyX port team of *classicthesis* is such a relief to the non-techie humanities researchers like me!

I thank Io Salmons, Myriam Cantú, Jingtao Zhu, Shenai Hu from the Acquisition and Pathology Lab at the UAB for their valuable suggestions and feedback throughout my doctoral work. I thank Isabel Crespí from the Centre de Lingüística Teòrica, UAB, who translated the abstract of my thesis into Catalan. I also thank Cristina Puigdollers at the CLT, UAB, for all the invaluable help with the documentation that helped me to submit my thesis on time.

I am indebted to my family, especially my mother, my sister and my little brother who stood by me when I wanted to give up a highly paid job in the States to pursue my passion for language acquisition research. Vishnu and Anu, thank you for being my most sincere critics and support. I would like to thank my partner Naru; his Buddha-like nature had helped me in every way to push forward when I wanted to give up.

It was a difficult journey, I was painfully slow, but I did this for myself, and I am happy I made it to the end through episodes of depression, procrastination, complete numbness and then adamantly pursuing my passion.

## DECLARATION

---

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

*Barcelona, July 2016*

---

Maya Leela



## ACRONYMS

---

ACC Accusative

ADVL Adverbial Suffix

CI Conceptual Intentional Interface

COMP Complementizer

CONJ Conjunction

DAT Dative

EMPH Emphasis

ERG Ergative

F Feminine

FUT Future tense

GEN Genitive

INF Infinitive

INST Instrumental

LOC Locative

M Masculine

NOM Nominative

PAST Simple Past Tense

PERF Perfective Aspect

POSS Possessive

PRES Present

P&P Principles and Parameters

PROG Progressive

QP Quotative Particle

RP Relative Participle

RM Relativized Minimality

SM Sensory Motor Interface

SMT Strong Minimalist Theory

SOC Sociative

UG Universal Grammar

VEPS Very Early Parameter Setting

## CONTENTS

---

### I INTRODUCTION

- 1 LANGUAGE ACQUISITION 2
  - 1.1 Parameters and Word Order Acquisition 6
    - 1.1.1 The OV/VO order 9
    - 1.1.2 Scrambling 14
  - 1.2 Rationale and Aims 17
    - 1.2.1 Aims 18
  - 1.3 Malayalam and Hindi-Urdu 18

### II THE SYNTAX AND FREQUENCY OF WORD ORDER

- 2 THE SYNTAX OF WORD ORDER IN HINDI-URDU AND MALAY-  
ALAM 22
  - 2.1 Hindi-Urdu 22
    - 2.1.1 Agreement, Tense and Case 22
    - 2.1.2 Word Order and Information Structure 25
  - 2.2 Malayalam 29
    - 2.2.1 Agreement, Tense and Case 29
    - 2.2.2 Word Order and Information Structure 31
- 3 WORD ORDER FREQUENCY ANALYSIS 42
  - 3.1 The Hindi-Urdu Spontaneous Speech Analysis 43
    - 3.1.1 Methodology 43
    - 3.1.2 Results 43
    - 3.1.3 The Child-Directed Speech corpus 46
  - 3.2 Malayalam Spontaneous Speech Analysis 49
    - 3.2.1 Methodology 49
    - 3.2.2 Results 49
  - 3.3 Discussion 51

### III THE ACQUISITION OF HINDI-URDU WORD ORDER

- 4 THE ACQUISITION OF THE OV ORDER 54
  - 4.1 Introduction 54
  - 4.2 Word Order Acquisition Theories 54

4.2.1	Word order acquisition by lexical learning	55
4.2.2	Word order acquisition guided by innate knowledge	60
4.3	The Acquisition of VO Order	72
4.4	The Acquisition Experiment of OV order in Hindi-Urdu	76
4.4.1	Rationale	76
4.4.2	Experimental Design	76
4.4.3	Results	83
4.5	Discussion and Conclusion	87
4.5.1	Usage based or Universal grammar?	91
 <b>IV THE ACQUISITION OF MALAYALAM WORD ORDER</b>		
5	THE ACQUISITION OF SCRAMBLING	96
5.1	Introduction	96
5.2	The Literature on the Acquisition of Scrambling	96
5.3	Experiment I	108
5.3.1	Rationale	108
5.3.2	Experimental Design	108
5.3.3	Results	113
5.3.4	Discussion	122
5.4	Experiment II	126
5.4.1	Rationale	126
5.4.2	Experimental Design	126
5.4.3	Results	130
5.4.4	Discussion	132
5.5	Experiment III	134
5.5.1	Rationale	140
5.5.2	Experimental Design	140
5.5.3	Results	144
5.5.4	Discussion	146
5.6	General Discussion	148
5.6.1	Overview	148
5.6.2	Relativized Minimality and the Malayalam Child Grammar	150

**V CONCLUSION**

**6 CONCLUSION AND FUTURE RESEARCH 153**

**VI APPENDIX**

**A APPENDIX 157**

**A.1 Malayalam Experiment I Test Sentences 157**

**A.2 Malayalam Experiment II Test Sentences 159**

**A.3 Malayalam Experiment III Test Sentences 161**

**VII BIBLIOGRAPHY**

**BIBLIOGRAPHY 165**

## LIST OF FIGURES

---

Figure 1	Parameter Hierarchy, Baker (2002)	8
Figure 2	Distribution of OV, VO and varying orders in the languages of the world (WALS)	9
Figure 3	Word order/Linearization parameter hierarchy, Biberauer and Roberts (2012)	12
Figure 4	The Antisymmetric approach of linear order	14
Figure 5	Malayalam sentence structure (Mohanani, 1982)	35
Figure 6	Derivation of SOV (Jayaseelan, 2001)	36
Figure 7	Derivation of OSV	37
Figure 8	Derivation of SVO	39
Figure 9	Derivation of OVS	40
Figure 10	Frequency of verb-initial and verb-final orders	46
Figure 11	Percentage of verb-initial and verb-final orders	48
Figure 12	Percentage of frequency of all word orders in Malayalam	51
Figure 13	Video clips depicting the causative and non causative action in the French Experiment	73
Figure 14	Proportion of mean looking time during the grammatical and ungrammatical condition (Franck et al., 2013)	74
Figure 15	Chulbuli cartoon clip	79
Figure 16	The character introduction video clip	80
Figure 17	The character identification video clip	80
Figure 18	The video clip introducing the novel action	81
Figure 19	An example from the ungrammatical sentence video sequence ( <i>chona</i> )	81
Figure 20	An example from the grammatical sentence video sequence ( <i>khalayaa</i> )	82
Figure 21	The heat-map for grammatical sentences (C= Causative, R= Reflexive)	83

Figure 22	The heat-map for ungrammatical sentences (C= Causative, R= Reflexive)	84
Figure 23	Proportion of looking time to causative action during the grammatical and ungrammatical sentences	86
Figure 24	Comparison between French and Hindi-Urdu results	90
Figure 25	A sample of the picture used in Experiment I	110
Figure 26	Results for Group I (mean age - 3;5 years)	115
Figure 27	Results for Group II (mean age - 4;3years)	116
Figure 28	Results for Group III (mean age - 5;3years)	117
Figure 29	The percentage of all responses in all the word orders for all the age groups	118
Figure 30	Results for the verbs used in SOV and OSV orders	119
Figure 31	Results for the verbs used in SVO and OVS orders	120
Figure 32	The experimental setup	129
Figure 33	Results	131
Figure 34	Results of Experiment II	132
Figure 35	The context for the verb pet/massage	134
Figure 36	Results of Experiment III	146

## LIST OF TABLES

---

Table 1	Raw score and percentage of different word orders in Hindi-Urdu	44
Table 2	Raw score and percentage of all word orders in child-directed speech	47
Table 3	Frequency of null arguments in Malayalam	50
Table 4	Raw score and percentage of different word orders in Malayalam	50
Table 5	Mean looking time in ms to the causative and reflexive clips during the grammatical and ungrammatical condition	85

Table 6	Chance level performance in causative action	86
Table 7	The comparative analysis between proportions of looking to the causative action during the grammatical and ungrammatical sentences	87
Table 8	Correct responses by experimental and control group (Otsu, 1994)	100
Table 9	Percentage of correct response (Murasugi and Kawamura, 2005)	103
Table 10	Subjects	109
Table 11	Verbs used in Experiment I	111
Table 12	Raw score and percentage of correct and incorrect responses	114
Table 13	Percentage and raw score of the verbs used in SOV and OSV orders	119
Table 14	Percentage and raw score of the verbs used in SVO and OVS orders	120
Table 15	Probability of correct answers for Group I (mean age 3.5 years)	121
Table 16	Probability of correct answers for Group II (mean age 4.3 years)	122
Table 17	Probability of correct answers for Group III (mean age 5.3 years)	122
Table 18	Verbs used in Experiment II	127
Table 19	Raw score and percentage of correct and incorrect answers	130
Table 20	Raw score and percentage of correct and incorrect answers per word order	131
Table 21	Verbs used in Experiment III	141
Table 22	Individual raw score per sentence type	145
Table 23	Raw score and percentage of correct answers	145



## Part I

### INTRODUCTION

## LANGUAGE ACQUISITION

---

Language acquisition research is directed towards answering a few fundamental questions. Questions like: is child grammar different from adult grammar, how different, and how and when does the child grammar evolve into the adult grammar. Many schools of thought attempt to answer these questions on acquisition. Classically they are divided into two approaches; namely the nativist and the emergentist approaches. The nativist theories (Chomsky, 1981) assume that children are born with innate language-specific knowledge, Universal Grammar (UG), and that language acquisition is predetermined by this knowledge. The emergentist (empiricist) theories (Tomasello, 1992; O Grady, 2010) propose that language is a learned behavior that is mastered through imitation of the external stimuli. While empiricism assumes language acquisition to be an item-based learning behavior, nativism proposes language acquisition to be the biological development of an autonomous species specific faculty of its own. Nativism is approached in detail here, with some reference to the empiricist proposals.

If a human child can acquire any language without any conscious effort or formal teaching, that should lead one to hypothesize about an innate biological capacity within the child that is ready for acquisition. This innate biological capacity should allow for the acquisition of all or any of the languages. The language acquisition literature within the generative framework has already provided ample evidence to the presence of such an innate biological capacity (De Villiers & Roeper, 2011). Though abstract grammar is proposed to be present in children from birth, language behavior of children exhibits a developmental pattern. If one assumes that abstract principles are innate, then the developmental course of language acquisition must be answered. One of the earlier accounts, the Continuity hypothesis (Pinker, 1984; Crain & Fodor, 1987), considered that all the principles of UG are fixed in child grammar, like in adult grammar, and the developmental course of acquisi-

tion is nothing but a transition through different grammars. According to the Continuity hypothesis, a construction in the child language that is ungrammatical according to the target language could be grammatical in another possible human language. The hypothesis proposes that all innate UG principles are functioning in children at very early ages, but children go through a transition of stages where they use constructions that are not found in their target grammar but found elsewhere in human language. From this standpoint, language acquisition can be viewed as language change (Crain & Pietroski, 2001), where a child explores different grammars within UG and finally settles to his/her target grammar/s in the immediate linguistic vicinity. This hypothesis endorses the nativist approach to language acquisition in such a way that it explains how errors in child grammar are constrained by UG. Though the hypothesis captures the errors that are seen in child grammar with respect to their native language, no explanation is given about the reason why some grammatical properties of the target language/s are not seen at some particular age. The Strong Continuity Hypothesis (Crain & Fodor, 1987) and the Weak Continuity Hypothesis (Clahsen, 1992) further instantiate the initial proposition of the Continuity hypothesis.

By criticizing the Continuity hypothesis, a later account, the Maturation hypothesis (Borer & Wexler, 1987, 1992) viewed the problem of learnability from a biological perspective. If the biological nature of language is taken into consideration, an important question arises. As Borer & Wexler (1987) ask, the question is:

*“why should the linguistic system be the only system which is fixed at birth and which does not mature?”* (Borer and Wexler 1987, p. 130)

It was proposed that not all UG principles are functioning at birth, but, as any other biological system matures, the UG principles also mature. The duration of the maturation of a principle accounted for the developmental course of language acquisition. Borer and Wexler claimed that the delay in the acquisition of A-chains is due to maturation. According to this proposal, maturation is constrained by the principles of UG. The maturation account answers why some linguistic structures or abilities emerge late in children.

From a pivotal standpoint that takes into account the Principles and Parameters framework, Culicover (1997) explains that both continuity and maturity could be accommodated. He proposed that, what is fixed at birth is competence and what is maturing is performance. Culicover claims that the competence is achieved by the setting of parameters at very early stages of life and the performance continues to mature over the biological growth. A related approach which explains the gradual development of grammar is the Modularity hypothesis (Fodor, 1983). Under this hypothesis, language is viewed as a domain specific, autonomous and innate mental module. The various domains of language are proposed to be interdependent and interactive with each other. According to Fodor, language acquisition is gradual because the rate of development of these domains and their interactions vary. The Modularity hypothesis is relevant to the topic under research in this thesis, word order, as recent Minimalist proposals (Chomsky, 1995, 1999, 2016) claim the linear order of constituents to be something that is specific to a particular interface domain within the language module.

#### *Word Order in Minimalism*

As Generative Grammar evolved into Minimalist Program, the proposals about the linear ordering of constituents changed. In the Strong Minimalist Theory (SMT) (Chomsky, 2005, 2013) it is claimed that at the computational level there is only Merge, and whatever is the result of Merge is devoid of order. Once the merged constituents reach the Sensory-Motor (SM) interface and only then, for externalization purposes, the rules of word order apply. In Chomsky's own words:

*"mapping to SM adds linear order, prosody, and detailed phonetic properties"* (Chomsky 2016, p. 18)

In this minimalist view, word order is an interface feature. To be specific, according to the SMT, linear order is a peripheral aspect of language. But there is no reason to assume that this property of SM does not fall into innateness. It is an established fact in neuroscience that the property of the nervous system to adhere to linear order is an inherent human capacity. The presence of linear order in SM, according to Chomsky,

follows from this physiological property of the sensory modality of the SM interface. And thus this should also be an innate capacity of the language faculty.

Supposing that innateness only held valid for the core recursive property of language (Hauser et al., 2002), then the question of how a young infant can parse word order effectively without formal training arises. It was shown that the English speaking children at two-word production stage are sensitive to the word order of their native language (Brown, 1973). Such empirical evidence is in accordance with the proposal that the knowledge of word order rests on innate mechanisms. That brings the discussion to the underlying assumption of this thesis: if knowledge of abstract word order is innate, then the early language behavior of children who acquire Hindi-Urdu and Malayalam would exhibit this knowledge.

In addition, there is another proposal that is evaluated in this thesis by virtue of the Strong Minimalist Theory. Chomsky (1995) proposed that linguistic rules are structure dependent and not linear order dependent. Given a choice between a minimal linear distance and a minimal structural distance for computation, the system chooses the structural distance computation. Chomsky (2016) gives examples like (1, 2) as evidence for this claim.

- (1) Instinctively, [birds that fly] [swim]].
- (2) Instinctively, [[the desire to fly] [appeals [to children]]].

In these examples, even though the linearly closest verb to the adverb is *fly*, there is no ambiguity in the interpretation that the adverb *instinctively* modifies the verbs *swim* and *appeal* which are far in terms of linear distance. So the system chooses the structural distance even though the linear distance gives eligible candidates for interpretation. This contrast is unexpected, given the general assumption that linear distance is simple to compute, and the structural distance is complex to compute. The experiments in Malayalam, conducted as part of this thesis, investigated this contrast.

The results reported in this thesis are about the acquisition of some idiosyncratic properties of languages. Though this thesis does not enumerate all parameters which determine word order but deals only with a

subset, it is important to introduce the Principles and Parameters framework, as this is the theoretical approach on which the investigation reported in this thesis was built. Next section elaborates on the literature pertaining to parametric variation in the acquisition of word order, followed by a description of the two features of word order which were investigated.

Before going into the features of word order tested, a simple question needs to be restated, which serves as a prime justification for language acquisition research. What is in language acquisition research for linguistic theories? According to Rizzi (1993) there are three main contributions:

1. If a given linguistic principle has been proposed to belong to UG, then if it can be proven that children exhibit that principle by early ages then this can be considered as empirical evidence for UG.
2. Stages of development can be explained using parameter theory, which in effect holds the explanatory capacity of UG mechanisms for development.
3. The evidence which could be hard to find from adult grammar can be found from child grammar and this will in turn support UG.

With these justifications, language acquisition research, then augments linguistics research in general.

## 1.1 PARAMETERS AND WORD ORDER ACQUISITION

Parametric linguistics is considered a sub-field of biolinguistics (Gianollo et al., 2008), in which extensive research is carried out to explain language typology by means of parameters, and about the setting of those parameters. As Gibson & Wexler (1994) note, linguistic diversity can be explained by the setting of parametric values. Despite cross-linguistic variation and differences in various parametric values, it has been observed that language acquisition takes place in a uniform fashion following the same trend with respect to quantity, quality, and time of acquisition. Plato's problem with respect to language acquisition

was addressed best by Chomsky's Principles and Parameters (P&P) approach (Chomsky, 1981, 1995). The P&P approach provides answers to questions like how effectively language diversity can be explained and how a developing mind can acquire language with minimal exposure and no formal teaching. Universal Grammar contains principles which constrain languages in general and parameters which can take different values defining a possible human language. It is the combination of principles and parameters that bring forth the distinctive properties of grammar. The diversity of human languages is thus explained on the basis of parametric choices. Based on the exposure to the linguistic data, UG aids in setting the parametric values that match the target grammar. This proposal is maintained in the minimalist program (Chomsky, 2005).

It seems that young infants do not depend on their external experiences to make up for poverty of stimulus. A direct or indirect genetic means which is internalized in the infant is what guides language acquisition and fills the gap which poverty of stimulus poses. Thus a child's internal grammar is the result of an integration of the genetic endowment, which is UG, with the primary linguistic data (Lightfoot, 2006).

Cross-linguistic variation is not random; there are definite patterns and, at least for some of these, a theoretical explanation based on parametric variation has been proposed. There are different types of parameters and, based on how they influence other parameters, they are argued to be organized in a hierarchical fashion (Baker, 2002; Biberauer et al., 2009; Biberauer & Roberts, 2012; Sheehan, 2014). One parameter may entail an array of complex properties and a change in this one parameter could render wide changes within closely related or otherwise similar languages. Also, change in one parameter can predict the possible outcome of another set of associated parameters.

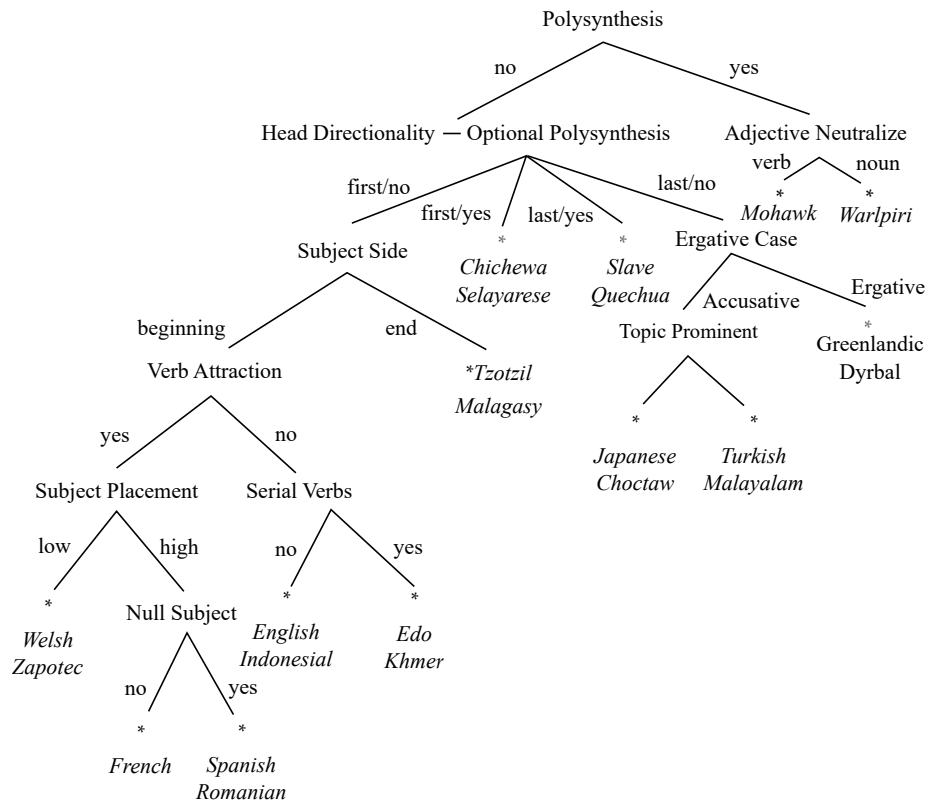
Baker (2002) has proposed a hierarchical representation of parameters which captures the parametric variation of possible human languages. If cross-linguistic variation is a result of different values taken by a finite number of parameters, then it should be possible to capture all the parameters in a single list. And any type of grammar should have a place in that parameter list. So parameters are ranked and listed on a

cause and effect basis, on how they affect each other and the power they employ on each other. Baker's ranking principle states that:

*"Parameter X ranks higher than parameter Y if and only if Y produces a difference in one type of language defined by X, but not in the other". (Baker 2002, p 163).*

Figure (1) gives the parameter hierarchy proposed by Baker (2002).

Figure 1: Parameter Hierarchy, Baker (2002)



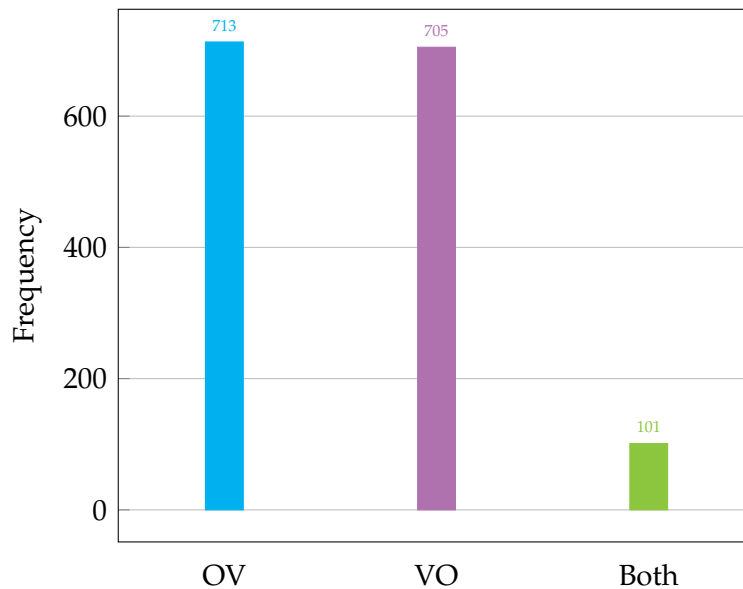
The parameter hierarchy by Baker is a list of macroparameters. In a single hierarchical structure, it covers all the proposed parameters. However, in parametric linguistics, the concept of sub-parameters has also been proposed. The later proposals put forth on the sub-parameters which determine the order of heads and complements are dealt with, in detail, in the next section.



### 1.1.1 The OV/VO order

Typological studies from the early sixties (Greenberg, 1963) identified the linear order of the object and the verb as one of the main loci of variation across languages. According to the data provided in WALS (Dryer & Haspelmath, 2013) out of the 1519 languages analyzed, 713 languages display OV order, 705 languages display VO order and 101 languages display varying verb-object orders (see Figure 2)

Figure 2: Distribution of OV, VO and varying orders in the languages of the world (WALS)



Languages like French and English exhibit consistent head-complement order. They are head-initial languages, and the cluster properties associated with the head-complement linear order are: the use of prepositions (example (3) in French), auxiliaries preceding main verbs (example (4) in French), nouns preceding complements (example (5) in French), and verbs preceding objects (example (6) in French).

- (3) *sur la table*  
on the table

'on the table'

- (4) *Susan a chauté une chanson.*  
 Susan has sung a song  
 'Susan has sung a song.'
- (5) *la fleur de cet arbre.*  
 the flower of that tree  
 'The flower of that tree.'
- (6) *Le chat chasse le rat.*  
 the cat chased the rat  
 'The cat chased the rat.'

In head-final languages like Hindi-Urdu and Japanese, the linear order displayed is complement-head. These languages present post-positions (example (7) in Hindi-Urdu), auxiliaries follow main verbs (example (8) in Hindi-Urdu), nouns follow complements (example (9) in Hindi-Urdu) and verbs follow objects (example (10) in Hindi-Urdu).

- (7) *mes ke upar.*  
 table POSS on  
 'on the table.'
- (8) *sita gaa rahi hein.*  
 sita sing PROG be-PRES  
 'Sita is singing a song.'
- (9) *us ped ke phool.*  
 that tree POSS flower  
 'The flower of that tree.'
- (10) *billi-ne choohe-ko bhagaayaa.*  
 cat-ERG rat-ACC chase-PERF  
 'The cat chased the rat.'

However, there are languages which do not adhere to this consistency. In Dutch and German not all phrases follow a harmonious ordering, VP is ordered as head-final, but PP and CP are ordered as head-initial. These languages fall into the head-variable category.

In early generative work, the assumption was that the directionality parameter is responsible for the canonical linear ordering of heads and complements in a language. Maintaining this assumption the head-final/head-initial parameter was proposed by Stowell (1981). Within  $v'$ , then, the order of Hindi-Urdu and Malayalam is the mirror image of the order in French and English. However, since there are languages which display both head-final and head-initial features, it became necessary to break down the parameters that govern linear ordering of heads and complements into further smaller parameters. Only in such an approach, the head-variable languages could be accommodated.

A different parameter hierarchy, developed by Biberauer & Roberts (2012) is in line with the above-mentioned assumptions. This new hierarchy was developed based on Baker's hierarchy, but with renewed features. Biberauer and Roberts produced a much more detailed description of parameters and their inter-dependencies. In this approach, parameters were divided into further sub-types. Macroparameters are the ones occupying the top most position in the hierarchy and as the algorithm goes down it gets split into Meso, Micro and Nano parameters. And this proposal provides an explanation about how the acquisition could take place. According to Biberauer and Roberts, UG is not predefined on parameter hierarchies but as a result of the interaction between primary linguistic data and UG it emerges as a by-product. The range of parametric variation is defined under four types of parameters.

“For a given value of  $v_i$  of a parametrically variant feature F:

- Macroparameters: all heads of the relevant type share  $v_i$
- Mesoparameters: all functional heads of a given category (eg: all verbal heads, all nominal heads, all theta bearing heads or all finite Cs) share  $v_i$
- Microparameters: a small subclass of functional heads (e.g. auxiliaries, pronouns) share  $v_i$
- Nanoparameters: one or more idiosyncratic lexical items are specified for  $v_i$ .”

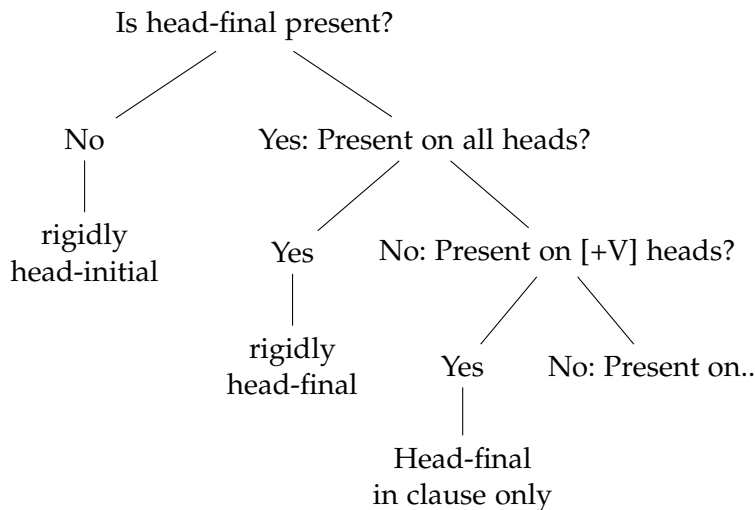
(Biberauer and Roberts 2012, p. 268)

Feature economy and input generalization are the two markedness conditions on which the hierarchies arise according to Biberauer & Roberts

(2012). The hierarchy is the result of a step by step process where children first apply a general rule to every category and only in the presence of counter-evidence they re-evaluate it.

The first few nodes of the parameter hierarchy for word order/Linearization proposed in Biberauer & Roberts (2012) is given in Figure 3. This hierarchy is particularly relevant as this thesis deals with the acquisition of word order. Though the term head-final is used to represent some nodes, the authors insist on viewing this as a cover term, which could either be a feature responsible for the movement of complements, or a Head Parameter at the PF.

Figure 3: Word order/Linearization parameter hierarchy, Biberauer and Roberts (2012)



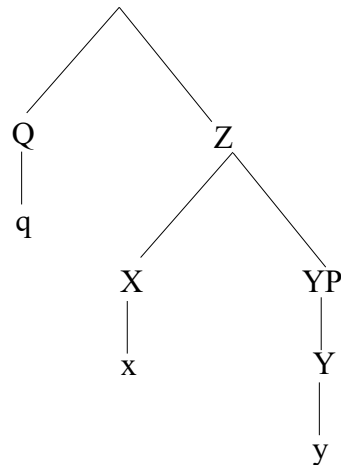
The higher node in this hierarchy is occupied by a macroparameter, and the effect of this parameter is seen on all head-complement order relations. In rigidly head-final and rigidly head-initial languages, all heads consistently follow the same pattern. The mesoparameter setting is explored only if the primary linguistic data presents evidence that only certain heads and complements follow the head-final order. The setting of a mesoparameter does not have as rigid an effect on grammar as the macroparameter. The last two sub-parameters, the micro and nano parameters, define the order for just a few heads or only the

behavior of some lexical items. Languages like Hindi-Urdu, Japanese and Malayalam would fall under the rigidly head-final branch in the hierarchy.

According to [Roberts \(2012\)](#), the evidence for setting macroparameters is ample in a child's linguistic environment and this enables the child to set the parameters very early. However, one cannot be sure about the time frame within which the setting of parameters travels down the hierarchy. Do children take longer to go from macro to meso parameter, for instance? The rate with which language acquisition takes place indicates otherwise. There may be a time difference, or there may not, and if there is a time difference, it may not even be a gap that can be tested through language performance at the exact time period of transition. The theory of syntactic cues ([Lightfoot, 2006](#)) and the theory of triggers ([Gibson & Wexler, 1994](#)) explain that very little amount of evidence from the external linguistic stimuli can set a parameter.

The directionality of heads and complements being explained as a product of a directionality parameter has received opposition. [Kayne \(1994\)](#), argued that within UG principles linear order reflects structural hierarchy. The universal order underlying language is the Spec-Head-Complement order. Kayne proposed the so-called antisymmetric approach, which explains the order of terminal nodes in a hierarchy to be the result of an asymmetric c-command relation of nonterminal nodes. For the terminal nodes ( $x\ y$ ), their surface order is the result of the c-commanding relation of  $X$  and  $Y$  (see Figure 4).

Figure 4: The Antisymmetric approach of linear order



Head-final and head-initial languages exhibit an overriding specifier-first order. Based on this symmetry between head-initial and head-final languages, the head-final order in the form of Spec-Complement-Head was proposed as a derived order via movement from a head-initial order in the Kaynian analysis, and this is assumed in this thesis.

### 1.1.2 *Scrambling*

Apart from the acquisition of the OV order, this thesis deals with the acquisition of scrambling. The term scrambling was introduced by Ross (1967) for variations in the linear ordering of NP and V. Ross defined scrambling as a stylistic variation rather than resulting from a syntactic rule. He argued that for free word order languages there is a scrambling universal. The universal scrambling rule would be that an NP and V can be arranged in any order and the language specific grammar would exert further restrictions. Ross provides examples from Latin and Russian, to substantiate these claims, where in Latin it is possible to place the adjective away from its noun, but in Russian this is not possible.

In the literature, there are two approaches with which scrambling is analyzed: one which assumes that there is no movement, and that scrambled orders are base generated (Mohan, 1982; Felix, 1992; Pierce, 2012) and another approach which assumes that scrambling is derived by movement (Jayaseelan, 1996; Kidwai, 1999; Fukui & Sakai, 2003; Mura-sugi & Kawamura, 2005; Manetta, 2012). Since Ross, the literature on scrambling has grown and now it is generally agreed that scrambling is a syntactic operation (Abels, 2015). Broadly the consensus is that scrambling as an umbrella term has to do with movement operations with implications in semantics-syntax or syntax-pragmatics. These views are corroborated based on facts like how scrambling affects either definiteness or specificity or how scrambling is used to code information structure. Researchers like Schaeffer (2000a) claim that scrambling is discourse-pragmatic; Avrutin & Brun (2001) explain that scrambling is the result of syntax-discourse interface activity. Saito (1989) explicitly argues that scrambling in Japanese is semantically vacuous, and that scrambling is undone in LF through radical reconstruction.

One of the variations in scrambling is about the type of scrambling which is allowed in a language. The categorization is based on whether scrambling is clause bound or not. Clause bound scrambling is termed local scrambling and is found in simple clauses. Long-distance scrambling, which is not clause bound, occurs outside of an embedded clause (Haider, 2000). Tada (1993) classifies scrambling in Japanese into Long scrambling, Middle scrambling and Short scrambling. Long scrambling (11) is not clause bound as middle-scrambling (12). Short scrambling is the altered order of indirect object and direct object (13). The examples are taken from Sugisaki & Isobe (2001).

- (11) *Mary-ni<sub>i</sub> Bill-ga [John-ga t<sub>i</sub> atta to] omotteiru.*  
 Mary-DAT<sub>i</sub> Bill-NOM John-NOM t<sub>i</sub> met that think  
 'Bill thinks that John met Mary.'

- (12) *Mary-ni<sub>i</sub> John-ga t<sub>i</sub> sono hon-o ageta.*  
 Mary-DAT<sub>i</sub> John-NOM t<sub>i</sub> that book-ACC gave  
 'John gave that book to Mary.'

- (13) *John-ga sono hon-o<sub>i</sub> Mary-ni t<sub>i</sub> ageta.*  
 John-NOM that book-ACC<sub>i</sub> Mary-DAT t<sub>i</sub> gave

'John gave that book to Mary.'

### *Relativized Minimality*

As a movement analysis of scrambling is assumed in this thesis, one of the proposals about how to constrain movement, Relativized Minimality (RM) (Rizzi, 1990), is considered. In simple terms, the theory of RM talks about constraints on movement by virtue of how one constituent can act as an intervener in the movement of another constituent. If there is a potential intervener between the target and its trace, then the sentence is rendered ungrammatical. However, intervention effects have varying degrees. A constituent becomes a potential intervener based on the morpho-syntactic features of the target and the intervener. And these features are the primary factors which determine the RM effect. There are three types of such relations listed out by Friedmann et al. (2009), which are: identity, inclusion, and disjunction. In an identity relation, both lexical items bear the same features and the sentence is rejected as ungrammatical in both child and adult grammar. In disjunction, intervener and target possess dissimilar features and the sentence is accepted by child and adult grammar. The relationship which is categorized as inclusion is relevant for the characterization of child grammar according to Friedmann et al. (2009). The structural configuration of an inclusion relation is shown below.

(14) +A, +B.....+A.....<+A,+B>

A and B represent sets of morpho-syntactic features. When the intervener possesses some features which the target also possess, then they fall into a subset relationship. Parsing of this subset relation requires a complex computing process. Adult grammar allows these sentences even if there is a potential intervener as adult grammar can handle sophisticated computing processes. But child grammar faces difficulties because of the requirement of complex computations. Among others, Friedmann et al. (2009) and Belletti & Rizzi (2013) describe the delay in the acquisition of object relative clauses using this approach. Features of the target and the intervener in relative clauses are in a subset-superset relation. The subject with a similar +NP feature as the object intervenes



between the object relative head and its trace (15). Due to this inclusion of features sentence analysis becomes difficult in child grammar and parsing of object relative clauses is delayed in child grammar.

(15) +R ,+NP (object relative head) .....+NP (subject) .....+R,+NP  
(object trace)

This has been argued to be at the source of the results in Modern Hebrew by [Friedmann et al. \(2009\)](#) and [Hu et al. \(2016\)](#) in Chinese, among many others. The investigation carried out in this thesis deals with the acquisition of scrambling in simple clauses and the RM theory is critically evaluated using the results.

## 1.2 RATIONALE AND AIMS

An important factor to be considered in language acquisition research is that, as [Crain & Thornton \(2000\)](#) suggest, language performance by children is not always the measure of their competence. But through adequate methodology, it should be possible to tap into the competence of the child, and that is the aim of this thesis. Following [Crain & Thornton \(2000\)](#), the investigations carried out were based on the assumption that with the appropriate methodology the knowledge of word order in child grammar can be tested at early ages and the competence measurements can be brought forth.

One of the primary hypothesis on which experiments were designed is Very Early Parameter Setting (VEPS) ([Wexler, 1998](#)). In the light of empirical evidence from early word order in the Germanic languages, Very Early Parameter Setting was proposed. Wexler formulates VEPS as:

*“basic parameters are set correctly at the earliest observable stages, that is, at least from the time that the child enters the two-word stage, around 18 months of age. The ‘basic parameters’ include at least the following:*

- a. Word order, e.g. VO versus OV (e.g. Swedish versus German)*
- b. V to I or not (e.g. French versus English)*

- c. *V2 or not (e.g. German versus French or English)*  
 d. *Null subject or not (e.g. Italian versus English or French)."*

(Wexler 1998, p. 29)

### 1.2.1 *Aims*

The primary aim of the thesis is to report empirical results about word order acquisition, from two less researched languages, Hindi-Urdu, and Malayalam. This is undertaken by investigating two distinct word order features of these languages:

1. To provide data on the acquisition of canonical unmarked order, OV, of Hindi-Urdu in native 19-month old children.
2. To provide data about the acquisition of marked scrambled word orders of Malayalam in native children within the age range of 2-5 years.

The first aim investigates the acquisition of a fixed feature in word order and the second aim investigates the knowledge of the flexibility feature of linear ordering.

As an introduction, a small description about the demographics of Hindi-Urdu and Malayalam is in order.

## 1.3 MALAYALAM AND HINDI-URDU

Malayalam and Hindi-Urdu are languages native to India. Though both languages are native to the same country, they belong to different language families. Each language has its own distinct script. Despite the fact that they are typologically unrelated, they share a few similarities in vocabulary and phonology as Sanskrit has influenced both these languages.

### *Malayalam*

Malayalam is a Dravidian language, which is the official language spoken in the state of Kerala in South India. It is one of the official lan-

guages of the constitution of India, and it has been given the status of a classic language. Malayalam is spoken by more than 38 million people who are spread across the globe. Records about both written and spoken Malayalam date back to the 12th Century AD. Malayalam is said to have originated from Tamil. In particular, the argument is made that Malayalam retained a feature of Tamil grammar for a period of time, but lost it subsequently as the language evolved. The feature is subject-verb agreement. It was present in old Malayalam (Varma, 2000). Diachronic evolution resulted in the elimination of this agreement feature and modern Malayalam does not have subject-verb agreement. During the ages when Sanskrit was flourishing, Malayalam was influenced greatly by Sanskrit. Literary Malayalam still retains a lot of Sanskrit vocabulary. It was after the contributions of Thunjath Ezhuthachan, who is referred to as the father of Malayalam, that the language achieved independent status from Tamil as well as from Sanskrit. Malayalam has its own script and an abundant resource of literature. There are many dialectal variations for Malayalam: one based on the geographical location, and the other based on religion. Generally, the dialects based on geographical location are: Northern, Central and Southern Malayalam. Nasrani and Mopla are the two dialects that are based on religion. The dialect Malabari, which is part of Northern Malayalam, presents the most noticeable differences from the standard literary language. The experiments carried out in this thesis used the dialect spoken in Southern Kerala, which is Southern Malayalam. Even though there are dialectal variations for spoken Malayalam, the literary language holds uniformity (Asher & Kumari, 1997).

### *Hindi-Urdu*

Hindi-Urdu is an Indo-Aryan language in the Indo-European family. Hindustani is the common name used to denote Hindi-Urdu. This language takes its origin from Khari Boli, a local language which was spoken in the subcontinent (Kachru, 1966, 2006). It was after the British colonization that Hindi and Urdu came to be considered two distinct languages. Hindi is written using the Devanagari script and Urdu is written using the Perso-Arabic script. The two varieties differ mainly in phonology and present minor differences in morphology (Mohanani,

1994). The vocabulary of Hindi is more influenced by Sanskrit and Urdu vocabulary is more influenced by Persian. Hindi-Urdu is consistently a verb-final language. Verb-initial sentences are used mainly in theatrical/poetic language. A form of poetry in Urdu, called *Shayari*, is abundant in verb-initial sentences, so is Hindi poetry. Hindi-Urdu spoken with a more Sanskritized form is found mainly in the Northern parts of India, and a more Persianised Hindi-Urdu is spoken in Pakistan. The division is more due to different political and religious reasons than linguistic reasons.

Though the linguistic research undertaken in both these languages is abundant, language acquisition research in generative grammar remains notably scarce, and to my knowledge reduces to Grebenyova (2012) and Narasimhan & Dimroth (2008). As mentioned earlier this thesis attempts to add data about the acquisition of Hindi-Urdu and Malayalam grammar to the acquisition literature in generative grammar.

The thesis proceeds as follows, through five more parts. Part 2 contains a description of the syntax of Hindi-Urdu and Malayalam word order, and a word order frequency analysis carried out for these two languages. In Part 3 the experiment conducted on Hindi-Urdu infants about the acquisition of OV order is reported and in Part 4 the experiments conducted in Malayalam speaking children about the acquisition of scrambling and reconstruction are reported. Part 5 comprises of general discussion and conclusion of the investigations reported in the thesis.

## Part II

# THE SYNTAX AND FREQUENCY OF WORD ORDER

## THE SYNTAX OF WORD ORDER IN HINDI-URDU AND MALAYALAM

---

To derive an accurate conclusion about what is it that the children know when they display a particular language behavior, it is important to look at the linguistic stimuli that elicited that behavior. In this section, features of adult grammar with respect to the word order in Hindi-Urdu and Malayalam are given. How different grammatical properties were considered in choosing the test sentences, and in constructing the pseudo-verbs are also described in this part.

Hindi-Urdu and Malayalam are verb-final languages, SOV is the unmarked word order in both. They exhibit the property of scrambling, with a high degree of freedom in the linear ordering of arguments, predicates and adjuncts. The marked word orders depict information structure variations in the grammar of these two languages. One of the main differences between the two with respect to freedom of word order is that Long distance scrambling is allowed in Hindi-Urdu, whereas in Malayalam it is not.

### 2.1 HINDI-URDU

The word order acquisition experiment conducted in Hindi-Urdu employed active transitive sentences in the SVO and VSO word orders with pseudo-verbs. The relevant syntax, morphology, and phonology of Hindi-Urdu with respect to these two orders and the descriptive grammar rules on the verb formation are detailed in this section.

#### 2.1.1 *Agreement, Tense and Case*

As the constructions used in the experiment included an agent and patient argument, and a predicate, details of the descriptive grammar with

respect to the case marking on nominals, subject-verb agreement, and tense marking on verbs are described in this section.

When the subject is not accompanied by a post-position, verbs agree with subject in person, number, and gender. The general pattern of agreement in simple clauses, when the subject is prohibited by a post-position for subject-verb agreement, is as follows: the main verb and the auxiliary (if there is any) agree with the direct object of the clause on number and gender as exemplified below.

- (16) *raam-ne roti khaayi.*  
 raam-M-ERG bread-F eat-F-PERF  
 'Ram ate (the) bread.'

- (17) *raam-ne kela khaaya.*  
 raam-M-ERG banana-M eat-M-PERF  
 'Ram ate (the) banana.'

In terms of tense marking, the simple past and future tenses are marked on the root verb (18, 19), and all other tenses are marked on auxiliary verbs (20) (Kachru, 1966).

- (18) *raam aaya.*  
 ram come-PERF  
 'Ram came.'

- (19) *raam aayega.*  
 raam come-FUT  
 'Ram will come.'

- (20) *raam aa rahaa hein.*  
 raam come PROG be-PRES  
 'Ram is coming.'

To indicate the simple past tense, the perfective participle, *-a*, is used (21). As exemplified, auxiliary verbs are not associated with simple past, or in other words, a verb with the perfective participle, when not accompanied with an auxiliary that indicates tense, marks simple past tense.

- (21) *raam-ne ek giit gaaya.*  
 raam-ERG one song sing-PERF  
 'Ram sang a song.'

The verbal paradigm of a verb in simple past tense is formed by adding *-a* to the end of the root verb, for singular masculine and when the verb ends in a consonant (22).

(22) Root verb *chal*

*chala* in singular masculine form

*chale* in plural masculine form

*chali* in singular feminine form

*chali* in plural feminine form.

When a root verb ends in a vowel, *-y* appears in addition to the *-a*, in between the final vowel of the verb and the perfect participle (23).

(23) Root verb *ga*

*gaya* in singular masculine form

*gaye* in plural masculine form

*gayi* in singular feminine form

*gayi* in plural feminine form

Regarding the case marking of Hindi-Urdu, the literature is not conclusive about the relation between transitivity and the choice between ergative and nominative case marking. According to Mohanan (1994), there are three types of verbs in Hindi-Urdu based on which case the subjects get assigned. The first type of verbs take subjects with only nominative case; the second type takes subjects with only ergative case, and the third type takes either nominative or ergative subjects. When a verb is perfective marked, the subject is marked with the ergative case (21).

The case marking on object nominals is dependent on the animacy of the object nominal and the definiteness of the object nominal. The objects in Hindi-Urdu receive the accusative marker *-ko* if they are animate, and inanimate objects receive the nominative null marker. Definite inanimate objects can take the accusative *-ko* marker as an exception.



## 2.1.2 Word Order and Information Structure

As mentioned earlier, Hindi-Urdu is a head-final language and is strictly post-positional. The unmarked word order in a ditransitive sentence is SUB-IO-DO-V (24) (Mahajan, 1990; Mohanan, 1994; Kidwai, 1999).

- (24) *mohan-ne kumar-ko paisa diya.*  
 mohan-ERG kumar-ACC money give-PERF  
 'Mohan gave money to Kumar.'

The canonical word order SOV corresponds to wide focus but the marked non-canonical word orders are always associated with non-neutral focus and varying information structure. The non-canonical orders are the result of scrambling of one or all of the following constituents: arguments (25, 26, 27, 28), and varying positions of V (29, 30, 31, 32) or adjuncts (33, 34).

- (25) IO S DO V  
*kumar-ko mohan-ne paisa diya.*  
 kumar-ACC mohan-ERG money give-PERF  
 'Mohan gave money to Kumar.'

- (26) DO S IO V  
*paisa mohan-ne kumar-ko diya.*  
 money mohan-ERG kumar-ACC give-PERF  
 'Mohan gave (the) money to Kumar.'

- (27) IO DO S V  
*kumar-ko paisa mohan-ne diya.*  
 kumar-ACC money mohan-ERG give-PERF  
 '(The) money was given to Kumar by Mohan.'

- (28) S DO IO V  
*mohan-ne paisa kumar-ko diya.*  
 mohan-ERG money kumar-ACC give-PERF  
 'Mohan gave (the) money to Kumar.'

- (29) V S IO DO

*diya mohan-ne kumar-ko paisa.*  
 give-PERF mohan-ERG kumar-ACC money  
 'Mohan gave (the) money to Kumar.'

## (30) S V IO DO

*mohan-ne diya kumar-ko paisa*  
 mohan-ERG give-PERF kumar-ACC money  
 'Mohan gave (the) money to Kumar.'

## (31) S IO V DO

*mohan-ne kumar-ko diya paisa.*  
 mohan-ERG kumar-ACC give-PERF money  
 'Mohan gave (the) money to Kumar.'

## (32) S DO V IO

*mohan-ne paisa diya kumar-ko.*  
 mohan-ERG money give-PERF kumar-ACC  
 'Mohan gave (the) money to Kumar.'

## (33) S IO DO V Adjunct

*mohan-ne kumar-ko paisa diya kitaab khareedne*  
 mohan-ERG kumar-ACC money give-PERF book buy-to  
*ke-liye.*  
 for  
 'Mohan gave the money to Kumar to buy (a) book.'

## (34) Adjunct S IO DO V

*kitaab khareedne ke-liye mohan-ne kumar-ko paisa*  
 book buy-to for mohan-ERG kumar-ACC money  
*diya.*  
 give-PERF  
 'To buy (the) book Mohan gave (the) money to Kumar.'

In addition to the exemplified orders, more permutations and combinations of arguments, predicates and adjuncts are possible. As the grammar is strictly head-final, auxiliaries follow the main verb (35).

- (35) *raam roti khaa rahaa hein.*  
 raam bread eat PROG be-PRES  
 'Ram is eating roti.'

Elements from an NP (36), PP (38), adjective phrase (40) and adverb phrase cannot be scrambled.

- (36) *anu-ne [sita-ka haar] kharida.*  
 anu-ERG sita-POSS necklace buy-PERF  
 'Anu bought Sita's necklace.'

- (37) \**sita-ka<sub>i</sub> anu-ne haar<sub>i</sub> kharida.*  
 sita-POSS anu-ERG necklace buy-PERF

- (38) *[kursi-ke upar] billi hein.*  
 chair-GEN on-top cat be-PRES  
 'The cat is on the chair.'

- (39) \**upar kursi-ke billi hein.*  
 on-top chair-GEN cat be-PRES

- (40) *anu-ne [bada ghar] kharida.*  
 anu-ERG big house buy-PERF  
 'Anu bought the big house.'

- (41) \**bada<sub>i</sub> anu-ne ghar<sub>i</sub> kharida.*  
 big anu-ERG house buy-PERF

Another feature of Hindi-Urdu is that it is a pro drop language where both the subject and object can be dropped (42).

- (42) *Kal dekhoongi.*  
 tomorrow see-FUT  
 '(I) will see (it/him/her/them) tomorrow.'

These are some of the general features of the grammar of Hindi-Urdu that are relevant to the investigation carried out in this thesis. The orders investigated in the acquisition experiment were SOV and VSO as mentioned in the beginning. The neutral-focused unmarked order SOV was used as the grammatical order, and it was contrasted with non-neutral-focused, marked VSO order as ungrammatical.

### Information Structure

Linear ordering of constituents has an important part in realizing information structure in Hindi-Urdu. The canonical word order marks neutral focus/wide focus, and it occurs as a declarative sentence, as an answer to the question *what happened?* Pre-verbal position indicates either definiteness or focus in marked word orders.

- (43) *raam-ne roti khaayi.*  
 raam-ERG bread eat-PERF  
 'Ram ate (the) bread.'

- (44) *roti raam-ne khaayi.*  
 bread raam-ERG eat-PERF  
 '(The) Bread, Ram ate.'

- (45) *khaayi raam-ne roti.*  
 eat-PERF raam-ERG bread  
 'Ate, Ram, (the) bread.'

Example (43) can only be an answer to the question *what happened?*, and the example (44), an OSV sentence, can only be an answer to *who ate the roti*, where the subject receives focus. In contrast, in example (45) the verb itself receives focus as in *Ram ate the bread*; it could also be a contrastive focus as in *Ram ate the bread* as opposed to *Ram threw the bread*.

- (46) *cheena hamse hamara maan, sammaan.*  
 snatch-PERF we-INST our pride honour  
 '(They) snatched our pride and honour from us.'

In the example (46), the verb receives a dramatic emphasis, this could be an answer to the question *what did they do to you?* When the happening itself receives a focus or contrastive focus or an emphatic stress, it is in the initial position.

The non-canonical word orders are always signaled by an altered stress and intonation and, according to Kidwai (1999), this forces the listener to divide the information into presupposed and asserted parts.

Then in a VSO sentence, the verb is asserted and has a specific intonation contour. The intonation contour for a neutral declarative sentence is different from that of a sentence with focus. Hindi-Urdu is categorized as an edge-prominence language, in which only boundary tones can change to convey pragmatic focus (Féry, 2010). In an SOV declarative sentence the intonation contour is as exemplified in (47).

(47) [[L\*H] [L\*H] [H\*L]]

In the non-canonical order VSO, the focused constituent retains a falling contour, but the post focused elements are lowered. So the intonation contour for VSO order is as in (48).

(48) [[H\*L] [L\*H] [L\*H]]

Accumulating all these facts it can be concluded that a question like *what happened?* can only accept a neutral canonical SOV answer and not a non-canonical VSO or OSV answer, for instance.

## 2.2 MALAYALAM

### 2.2.1 Agreement, Tense and Case

Malayalam is a nominative-accusative language, with bound suffixes and post-positional markers to indicate the syntactic and semantic functions ((49 - 51) where post-positions are given in capitals) (Asher & Kumari, 1997).

(49) *kathi KOND kai murichu.*  
knife WITH hand cut-PAST  
'Hand was cut with the knife.'

(50) *avanu VENDI jeevikkunnu.*  
he-DAT FOR live-PRES  
'Living for him.'

(51) *athine-KAAL valuth.*  
that-THAN big  
'bigger than that.'

The order of the post-positions, case markers, and bound suffixes in an NP is as follows: stem + (derivational suffix + plural marker) + case marker (52)

(52)

Verb stem - *kutti* ('child')

stem + plural marker - *kutti+kal* ('children')

stem + plural marker + case marker - *kutti+kal+e* (children-ACC)

The subject of a transitive verb, except experiencer verbs, is marked with an un-suffixed null nominative case. Animate direct objects and indirect objects are marked with an accusative marker *-e* (53). Inanimate objects are marked with the nominative null marker (54).

(53) *avan avale kandu.*  
he she-ACC saw-PAST  
'He saw her.'

(54) *avan pusthakam vaangi.*  
he book buy-PAST  
'He bought the book.'

The infinitive form of a verb root ends in *-uka*, and the transitive form is obtained by omitting the *-uka*, and adding the appropriate tense/aspect markers. Past tense is marked by adding the vowels *-i* or *-u* to the verb stem. The present continuous tense is marked on a verb stem using the suffix *-unnu* (55).

(55)

*odə* ('run') - verb stem

*od-uka* ('to run') - infinitive form

*od-i* ('ran') - past tense

*od-um* ('will run') - future tense

*od-unnu* ('running') - present continuous

According to the parameter hierarchy given by Biberauer et al. (2009) on null arguments, Malayalam is classified as radical pro-drop. The subjects, direct objects, and indirect objects can be dropped felicitous in an appropriate discourse context (56).

- (56) *collegeil vech kandappol koduthu.*  
 college-LOC at see-then give-PAST  
 '(I) gave (it) (to him/her) when (I) saw (him/her) at the college.'

### 2.2.2 Word Order and Information Structure

The canonical order of the verbs and arguments in a ditransitive Malayalam sentence is S-IO-DO-V (57a). Like Hindi-Urdu, Malayalam allows scrambling of arguments (57b), and verbs in various positions (57c) as well as adjuncts (57d - 57f).

#### (57) (a) S-IO-DO-V

*appu ammoonu aa pusthakam koduthu.*  
 appu ammu-ACC that book give-PAST  
 'Appu gave that book to Ammu.'

#### (b) DO-S-IO-V

*aa pusthakam appu ammoonu koduthu.*  
 that book appu ammoonu-ACC give-PAST  
 'Appu gave that book to Ammu.'

#### (c) V-S-IO-DO

*koduthu appu ammoonu aa pusthakam.*  
 give-PAST appu ammu-ACC that book  
 'Appu gave that book to Ammu.'

#### (d) *appu innale oru pusthakam vaangichu.* appu yesterday a book buy-PAST 'Appu bought a book yesterday.'

#### (e) *innale appu oru pusthakam vaangichu.* yesterday appu a book buy-PAST 'Appu bought a book yesterday.'

- (f) *appu oru pusthakam vaangichu innale.*  
 appu a book buy-PAST yesterday  
 'Appu bought a book yesterday.'

The verb initial order is uncommon in spoken language. It is found only in poetry and in theatrical language.

#### *Types of scrambling in Malayalam*

Among the different types of scrambling given by Sugisaki & Isobe (2001) and Tada (1993) for Japanese, the ones present in Malayalam are short scrambling and middle scrambling. Short scrambling is the term given for the altered order of the direct object and the indirect object within the VP. As exemplified in (58b) the canonical IO-DO (58a) is scrambled into DO-IO order.

- (58) (a) *appu sitaykku aa pusthakam koduthu.*  
 appu sita-DAT that book give-PAST  
 'Appu gave that book to Sita.'  
 (b) *appu aa pusthakam sitakku koduthu.*  
 appu that book sita-DAT gave  
 'Appu gave that book to Sita.'

Clause internal scrambling is termed as middle scrambling, where an object argument moves to the sentence initial position. DO-S-IO-V (59a) and IO-S-DO-V (59b) are two examples of middle scrambling from Malayalam.

- (59) (a) *aa pusthakam appu sitakk koduthu.*  
 that book appu sita-DAT gave  
 'Appu gave that book to Sita.'  
 (b) *sitakku appu aa pusthakam koduthu.*  
 sita-DAT appu that book gave  
 'Appu gave that book to Sita.'

Scrambling outside a clause boundary, long scrambling, is not allowed in Malayalam.



- (60) (a) *appu [amma sitaye kandennu] paranju.*  
 appu [mother sita-ACC saw-that] say-PAST  
 'Appu said that mother met Sita.'
- (b) \**sitaye appu amma kandennu paranju.*  
 sita-ACC appu mother saw-that say-PAST

#### *Restrictions on the Malayalam scrambling*

The freedom of word order is restricted within the clause boundary in certain conditions. Since the SOV order is semantically neutral and is used in declarative sentences, when the nominative and accusative markers are null, the first NP is taken as the subject and the second NP as the object (61 a, 61 b). The word order exhibits a freezing effect in such cases. In addition to the null case markers, when two arguments are marked with the same case markers, that also causes the word order to exhibit a freezing effect (Mohanani & Mohanani, 1994). The example (61 c) and (61 d) exemplify sentences with two consecutive arguments with dative marking, and examples (61 e) and (61 f) shows two consecutive arguments with accusative marking. Notice how agent and patient/theme change in the translation.

- (61) (a) *vellam neeraavi aayi.*  
 water steam be-PAST  
 'Water became steam.'
- (b) *neeraavi vellam aayi.*  
 steam water be-PAST  
 'Steam became water.'
- (c) *ammaykk achanu vishakkunna samayam*  
 mother-DAT father-DAT hunger-QP time  
*krithyamaayi ariyaam.*  
 correct-ADVL know-FUT  
 'Mother knows the exact time when father is hungry.'
- (d) *achanu ammaykk vishakkunna samayam*  
 father-DAT mother-DAT hunger-QP time  
*krithyamaayi ariyaam.*  
 correct-ADVL know-FUT

'Father knows the exact time when mother is hungry.'

- (e) *raaman sita-ye lakshmana-ne elppichu.*  
 raaman sita-ACC lakshmanan-ACC gave  
 'Raman gave Sita to Lakshmanan.'

- (f) *raaman lakshmana-ne sita-ye elppichu.*  
 raaman lakshmanan-ACC sita-ACC gave  
 'Raman gave Lakshmanan to Sita.'

The type of phrase also imposes a restriction on Malayalam scrambling. All the phrases strictly observe the head-final alignment; pre-nominal modifiers cannot scramble out of the NP (62 - 63), adjectives cannot scramble out of an AP (64 - 65), adverbs cannot scramble out of an AdvP (66 - 67), and post-positions cannot scramble out of the PP (68 - 69).

- (62) *unniyude aa<sub>i</sub> kuda<sub>i</sub>.*  
 unni-GEN that umbrella  
 'Unni's that umbrella.'

- (63) \**aa<sub>i</sub> unniyude kuda<sub>i</sub>.*  
 that unni-GEN umbrella

- (64) *enth sundaramaaya shabdham.*  
 what wonderful-is voice  
 'what a wonderful voice.'

- (65) \**enth shabdham sundaramaaya.*  
 what voice wonderful-is

- (66) *valare vegam odi.*  
 very fast run-PAST  
 'ran very fast.'

- (67) \**vegam valare odi.*  
 fast very run-PAST

- (68) *oonu kazhinjittu purappedaam.*  
 lunch finish-after leave-will  
 '(we/I) will leave after lunch.'
- (69) \**kazhinjitt oonu purappedaam.*  
 finish-after lunch leave-will

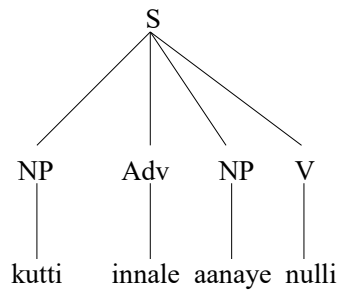
One of the main restrictions that freedom of word order faces has to do with information structure. The contexts in which the specific word orders face restriction for scrambling are explained in the next section.

#### 2.2.2.1 The Syntax of Malayalam scrambling

The earlier accounts (Mohanana, 1982) proposed a non-configurational flat structure to Malayalam word order. The scrambled sentences were treated as being freely base generated as given in Figure 5 for the example (70).

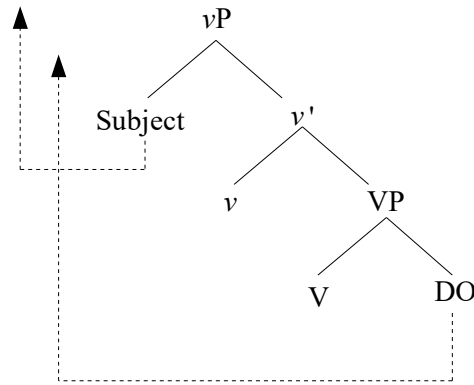
- (70) *kutti innale aanaye nulli*  
 child yesterday elephant-ACC pinc-PAST  
 'The child pinched the elephant yesterday.'

Figure 5: Malayalam sentence structure (Mohanana, 1982)



On the contrary, recent accounts (Jayaseelan, 1996) treat word order and alterations as the result of movement for topicalization or focusing. Following Kayne (1994), Jayaseelan (2001) assumes that the SOV order is derived from the universal Spec-Head-Complement order. So the basic word order, SOV, is a result of the object moving to the specifier position of a functional head, and the subject moving to the SPEC, IP as in Figure 6.

Figure 6: Derivation of SOV (Jayaseelan, 2001)



The SOV order is used in contexts which demand neutral focus. This condition is the same as that of Hindi-Urdu. The SOV order does not demand a presupposed knowledge from the listener about the context in which it appears. A question or a context that elicits wide focus results in the SOV order (Kidwai, 1999).

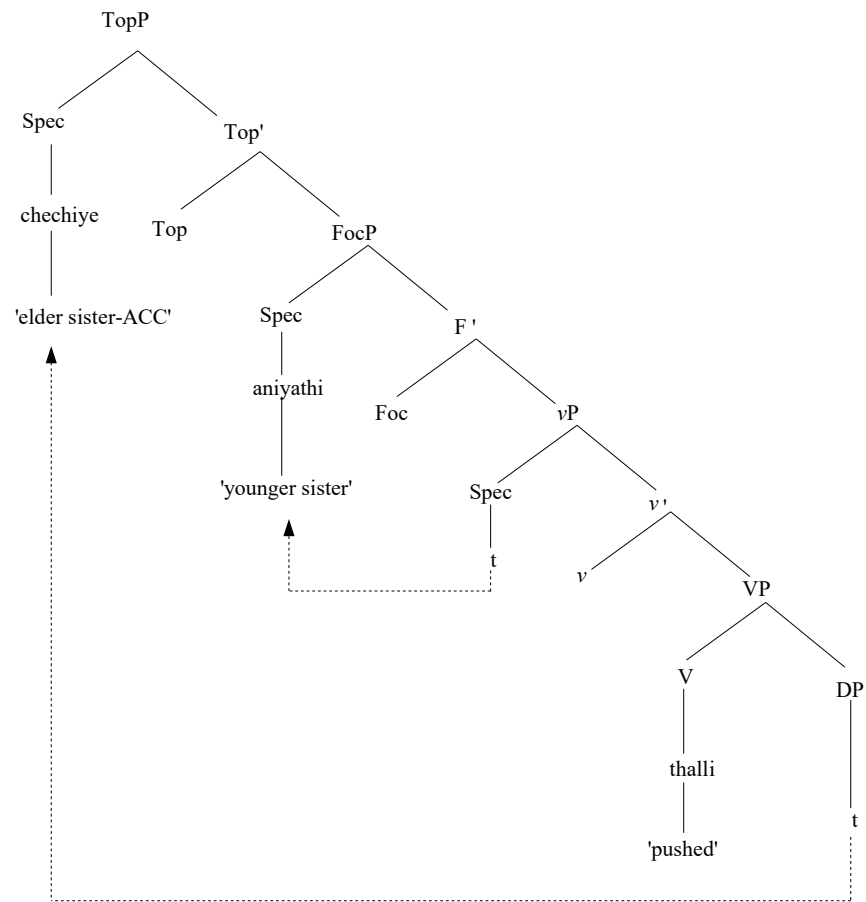
The marked word orders attest a difference in information structure. Among the non-canonical word orders in Malayalam, the ones that were tested in the acquisition experiments were OSV, OVS, and SVO. Adhering to the scope of this thesis, the syntax of these word orders are dealt with in detail.

In a discourse where the subject receives focus, the object moves to the specifier of the Topic Phrase and the subject moves to the specifier

of the Focus Phrase. The derivation of an OSV sentence (71) is as shown in Figure 7 if we follow Jayaseelan (2001).

- (71) *chechiye aniyathi thalli.*  
 elder sister-ACC younger sister push-PAST  
 'The younger sister pushed the elder sister.'

Figure 7: Derivation of OSV



The orders SVO and OVS contain post-verbal elements which may be analyzed as the result of the remnant movement of VP. Jayaseelan (1996)

claims that the arguments which appear to the right of V are said to be something of an afterthought. These post-verbal elements can be a DP (72), a PP (73), or adverbials (74).

- (72) *raamante oppam irangi sita.*  
 raaman-GEN with left sita  
 'Sita left with Raman.'

- (73) *hanuman kandu sitaye ashokavanathil vech.*  
 Hanuman saw sita-ACC ashoka-forest-LOC at  
 'Hanuman saw Sita at Ashoka forest.'

- (74) *sita poyikkaanum chilappo.*  
 sita went-may perhaps  
 'Perhaps Sita left.'

The order of the post-verbal elements seems to be the mirror image of the information structure order in the pre-verbal position. The arguments bearing focus or definiteness appear right adjacent to the verb and the topics or indefinite arguments appear at the right edge. The order of the topic and focus in pre-verbal position is Top-Foc-V (old-new-V), and in the post-verbal position the order is V-Foc-Top (V-new-old) (75 a - 75 d).

- (75) (a) V Sfocus Oold

*vaayichu RAAMAN aa ezhuth.*  
 read raaman-EMPH that letter  
 'Raman read that letter.'

- (b) V Ofocus Sold

*vaayichu aa ezhuth raaman.*  
 read-PAST that letter raaman  
 'Raman read that letter.'

- (c) V Onew Sold

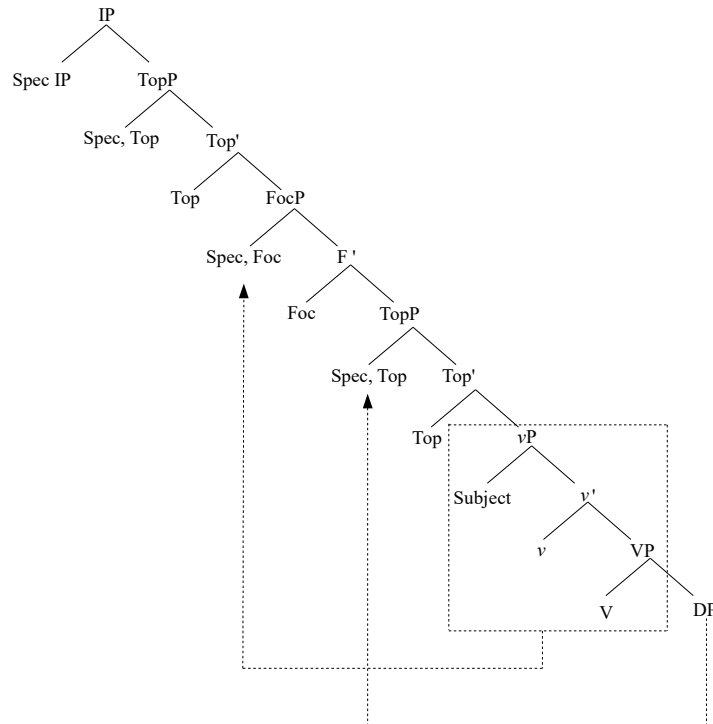
*kandu sitaye raaman.*  
 saw sita-ACC raman  
 'Raman saw Sita.'

(d) V Snew Oold

*kandu raaman sitaye.*  
 saw raaman sita-ACC  
 'Raman saw sita.'

According to Jayaseelan (2001), IP internal iterable TopP positions are the landing sites of scrambled arguments. As mentioned earlier, for a mono-transitive sentence, the argument appears in post-verbal position as a result of the remnant VP-movement. As Jayaseelan notes, in an SVO sentence, the subject and the verb move to SPEC,FP and F respectively. The object occupies the TopP below the FocP, which appears in post-verbal position in the surface order. The SVO sentence in the example (76) has the derivation as shown in Figure 8.

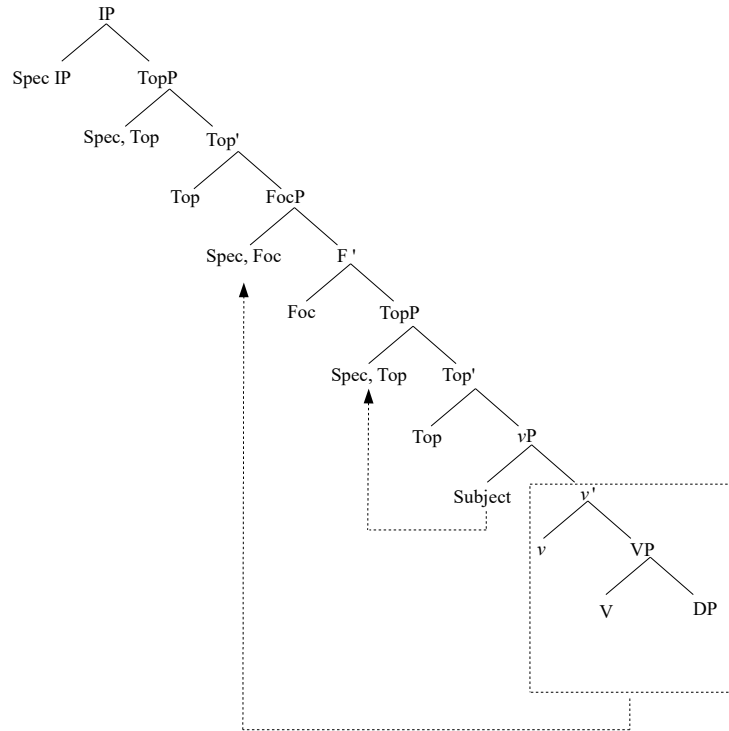
Figure 8: Derivation of SVO



- (76) *aniyathi thalli chechiye*  
 younger sister push-PAST younger sister-ACC  
 'The younger sister pushed the elder sister'.

In an OVS sentence, the object and the verb move to Spec,Foc, and the subject moves to Spec,Top and appears post-verbally in the surface order. The derivation of an OVS sentence (77) is given in Figure 9.

Figure 9: Derivation of OVS



- (77) *chechiye thalli aniyathi*  
 elder sister-ACC push-PAST younger sister  
 'The younger sister pushed the elder sister'.



To conclude; Malayalam scrambling is clause bound. Arguments cannot scramble outside an embedded clause, but they can scramble as a whole behaving like the argument of the matrix clause. Within the embedded clause the verb and the complementizer cannot scramble and remain on the right edge of the sentence. This restriction makes Malayalam scrambling similar to the scrambling found in Turkish (Batman-Ratyosyan & Stromswold, 1999), but different from Hindi-Urdu and Japanese. Malayalam scrambling involves A' movement.

WORD ORDER FREQUENCY ANALYSIS

---

Children acquiring languages with scrambling are exposed to a milieu of word order variations as their primary linguistic data. Acquisition of these languages includes both comprehension and production of all of the possible word orders in their corresponding information structure frame. So in order to investigate the acquisition of word order in languages with scrambling it was pertinent to obtain data on the frequency of occurrence of different word orders in spontaneous speech production. Also, data in such manner can be used to validate or dispute different theoretical claims. For instance, different claims about acquisition, understood as a process of imitative learning, cue-based learning or frequency-based learning, could be critically analyzed using a comparison of the data obtained from the word order acquisition experiments with the frequency of occurrence of the same word orders. If the results of the acquisition experiments reveal early acquisition of the less frequently occurring word orders, then that can be used in critically analyzing the frequency based accounts. In the same line, if the word orders which occur less frequently are found to be parsed by young children then this information would serve as evidence for early parameter setting. Additionally, it was essential to obtain the frequency of marked and unmarked word orders in spontaneous speech for two more reasons. First to get a clear picture about the occurrence of marked word orders in day-to-day life, and to compare and contrast this with the occurrence of unmarked word orders. Second, to determine which orders should be included in the acquisition investigation and to design the methodology accordingly.

There was no published data on the frequency of word order available for either of the languages investigated. Therefore the presence and frequency of different word orders in spontaneous speech production were analyzed as part of this thesis work. The analysis was done for both Hindi-Urdu and Malayalam spontaneous speech. The methodol-

ogy used for the analysis of Hindi-Urdu was different from the one used for Malayalam, as the structures tested were different in these two languages. Details about the differences in the methodology are explained in the following sections. The analysis and the results for both Hindi-Urdu and Malayalam are given in the section after that, and a comparison is presented at the end of the discussion section.

### 3.1 THE HINDI-URDU SPONTANEOUS SPEECH ANALYSIS

#### 3.1.1 *Methodology*

##### 3.1.1.1 *Design*

Conversations in the comment's section of various news articles in an e-paper called the *Navbharat Times* (NavbharatTimes.com/articles) were used as data source. A total of 5200 active sentences were included in the analysis, i.e. only passive sentences were excluded. Both ditransitive and monotransitive sentences were counted in for the analysis. The sentential arguments and one-word arguments were counted separately.

##### 3.1.1.2 *Coding*

Each sentence was entered into its respective word order category in an excel sheet, and when the same word order was encountered again, it was counted into the same category. The sum of each word order from each category and the total sum were calculated at the end.

#### 3.1.2 *Results*

It was found that the predominant order is SOV in Hindi-Urdu spontaneous speech. It occurred for 44% from the total number of sentences, and verb final orders were 73.6% (see Table 1).

Table 1: Raw score and percentage of different word orders in Hindi-Urdu

Word order	Count	Percentage
SOV	2264	43.5%
OV	566	10.9%
SV	449	8.6%
OSV	313	6%
SVO	286	5.5%
OVS	180	3.5%
SVOc	124	2.4%
VO	120	2.3%
S DO V IO	108	2.1%
Adv V	102	2%
DO S V IO	89	1.7%
NV: No Verb	79	1.5%
DO IO S V	74	1.4%
DO V IO	67	1.3%
V	74	1.4%
VS	53	1.2%
DO IO V	52	1%
VOS	50	1%
VSO	50	1%
DO S IO V	40	0.8%
VSOc	17	0.32%
VOc	15	0.3%
IO DO V	12	0.23%
IO S V DOc	9	0.2%
IO V DO	7	0.13%
Total	5200	

From the total sentences, 56% were of non-canonical orders. Some examples of the scrambled word orders found from the corpus are given

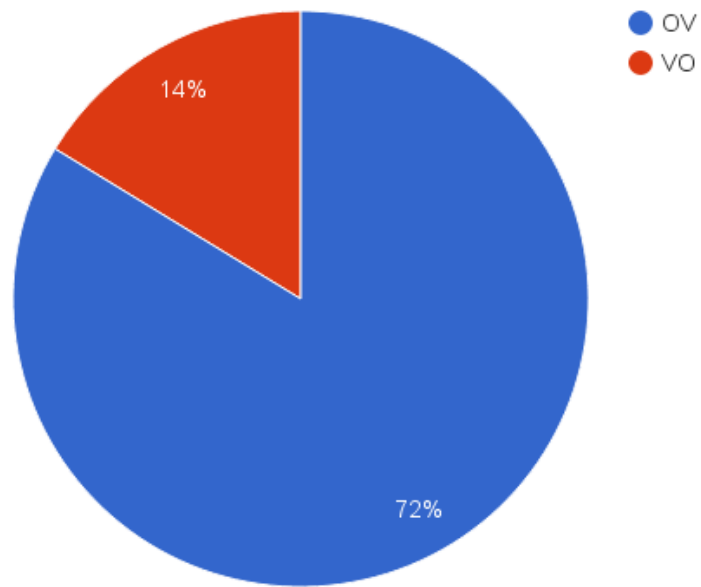
below, the example (78) is a verb initial sentence, and the example (79) is verb medial.

(78) *lage raho pyare.*  
 keep doing (it) beloveds  
 'Keep doing it, you beloved people.'

(79) *phir-to ab koyi nahi rok sakta modi-ji-ko.*  
 then now none not stop possible Modi-Honour-ACC  
 'Then no one can stop Modi now.'

Out of all the sentences, 85% (4443) were with an overt object. From the sentences with an overt object, OV order was present for 72% (3197); VO order was found only for 14% (611). Thus, more than 70% of the sentences containing an overt object displayed the OV order (Figure 10).

Figure 10: Frequency of verb-initial and verb-final orders



### 3.1.3 *The Child-Directed Speech corpus*

In addition to the adult spontaneous speech data, a small sample of 500 sentences were analyzed from child-directed speech. This was done in order to confirm that the word orders found in the adult spontaneous speech are attested in child-directed speech as well. Though the sample size is not enough for a proper comparison, a rough picture of the frequency of different word orders was drawn. The samples were homemade video recordings sent by mothers, on request, for the analysis and a few videos found on YouTube. Coding was done in a similar

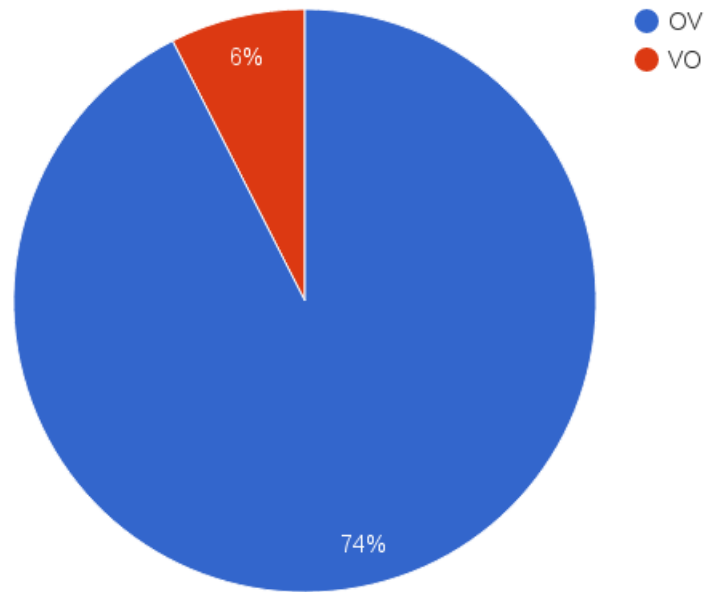
manner as the adult corpus analysis, except sentential arguments were not counted separately. All active sentences were included in the analysis. No passive sentences were found in the data. The results showed that 42% of the sentences were of SOV order (see Table 2 ).

Table 2: Raw score and percentage of all word orders in child-directed speech

Word order	Count	Percentage
SOV	210	42%
OSV	79	16%
OV	32	6.4%
SV	40	8%
Adj OV	1	0.2%
Adv V	10	2%
VO	5	1%
VSO	8	1.6%
VOS	8	1.6%
VS	10	2%
SVO	13	2.6%
OVS	80	16%
Adj VS	1	0.2%
Adj S	1	0.2%
Adj S	2	0.4%
Total	500	

From the 500 sentences, a total of 436 sentences (87%) contained an overt object. Among this, the OV order was found for 74% and, the VO order for 5% (see Figure 11).

Figure 11: Percentage of verb-initial and verb-final orders



In the adult spontaneous speech and in the child-directed speech, the percentage of OV order is very similar, i.e., more than 70% of the sentences were of OV order. The adult corpus showed 72% of OV order, and the child-directed speech showed 74% of sentences in the OV order. The VO order was found in 6% of the sentences in the child-directed speech, and 14% in the adult spontaneous speech.

The frequency analysis demonstrates that both OV and VO orders are attested in the Hindi-Urdu spontaneous speech. The frequency is considerably different among the two orders.



### 3.2 MALAYALAM SPONTANEOUS SPEECH ANALYSIS

#### 3.2.1 Methodology

Speech samples were collected from three types of discourse: interviews, discussions or debates, and conversations. Interviews were taken from the *Mathrubhumi News* portal website. Discussions, debates, and conversations were taken from various posts on Facebook. All declarative sentences were included in the counting, but clefts, passives, and ditransitive sentences were excluded. A total of 5598 monotransitive sentences were counted. Only monotransitive sentences were included because the acquisition experiments employed only these sentences. The sentential embedded clauses were counted as a single object or subject respectively. The sentences with null arguments were also included in the count. The coding was done in the same way as Hindi-Urdu analysis (3.1.1.2). Some examples of the marked word orders from the data are given below: OSV (80), OVS (81).

- (80) *kaaranam ayyappanu njan aarokkeyo aanu.*  
 because ayyappan-DAT I somebody is  
 'Because I am someone for Ayyappan.'
- (81) *ethrayo kaalamaayi theateril poyi cinema*  
 how-long time-is theater-LOC go-PAST cinema  
*kaanaarillaayirunnu njan.*  
 see-no-is-PastProg I  
 'I hadn't been going to watch movies in theater for so long.'

#### 3.2.2 Results

The least occurring order found in the spontaneous speech was verb-initial, and the most frequently occurring order was the unmarked SOV order. Significant amount of sentences with null arguments were also found, results of which are detailed in Table 3.

Table 3: Frequency of null arguments in Malayalam

	Raw frequency	Percentage
Sentences with Null arguments	1835	33%

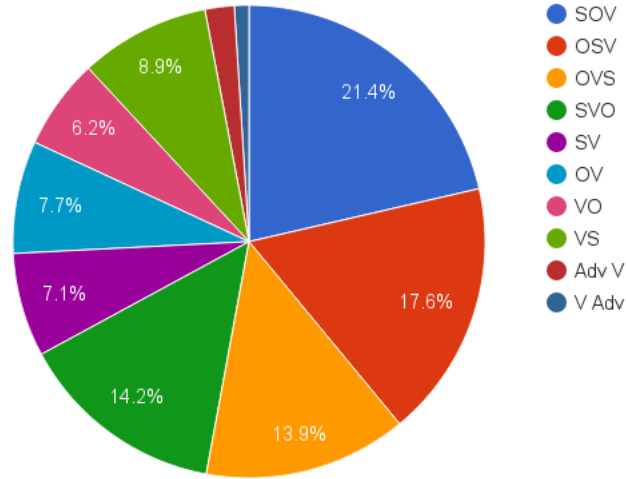
The different word orders found were SOV, OSV, OVS, SVO, SV, OV, VO, VS, Adverb V and V Adverb. The count and percentage of frequency all the orders from the data are given in Table 4.

Table 4: Raw score and percentage of different word orders in Malayalam

Word Order	Count	Percentage
SOV	1200	21.40%
OSV	985	17.60%
OVS	780	13.90%
SVO	798	14.20%
SV	400	7.10%
OV	432	7.70%
VO	345	6.20%
VS	498	8.90%
Adv V	89	2%
V Adv	71	1%
Total	5598	

From all the sentences analyzed, 55% (3106) of the sentences were verb-final. So the verb-final unmarked order is attested for more than 50% of the sentences in the spontaneous speech production. The most frequently occurring word orders were SOV, OSV, OVS and SVO ( see Figure 12).

Figure 12: Percentage of frequency of all word orders in Malayalam



### 3.3 DISCUSSION

The frequency of word orders in Hindi-Urdu and Malayalam spontaneous speech production were calculated. The results revealed that both marked and unmarked word orders are attested in the spontaneous speech. The most frequently occurring word order was the unmarked verb-final order. In the Malayalam data analysis, ditransitive sentences were excluded unlike in the Hindi-Urdu analysis. The experiment conducted in Hindi-Urdu was about the acquisition of OV order, and hence the transitivity of the sentences did not matter, but it was only necessary to calculate the order of object and verb. In that respect, the spontaneous speech analysis of Malayalam does not give a full picture of the word order of all transitive sentences in the language, but only the frequency of the word orders in monotransitive sentences. However, this analysis

served the purpose of determining which orders were more frequent in monotransitive sentences so that those orders could be chosen for the experiment. Also, since the methodology for Malayalam word order acquisition in this thesis only accommodated monotransitive sentences, adhering to analyzing the frequency of only those sentences was suitable. A child-directed speech corpus was analyzed only for Hindi-Urdu, for Malayalam no child-directed speech corpus was available, and gathering data was not possible due to time constraints.

Out of the 4398 monotransitive sentences in Hindi-Urdu, 3503 sentences, i.e., 80%, were verb-final. In Malayalam, however, only 55% of the sentences were verb-final. For the acquisition of the head-final (OV) order experiment this high percentage of verb-final sentences in spontaneous speech production aided in designing the methodology. Malayalam was found to have more diversity than Hindi-Urdu with respect to the presence of marked word orders. While Malayalam had 33% of sentences with null arguments, Hindi-Urdu had only 28% of sentences with null arguments, but this difference might disappear with a bigger corpus.

Part III

THE ACQUISITION OF HINDI-URDU WORD  
ORDER

## THE ACQUISITION OF THE OV ORDER

---

### 4.1 INTRODUCTION

Young infants acquiring language must have adequate knowledge of the surface word order of their native language in order to parse grammatical sentences. Studies conducted on the acquisition of word order has proven that children exhibit the knowledge of the word order of their native language from early stages of speech production (by one-word production stage) (Brown, 1973; Pinker, 1995; Hirsh-Pasek & Golinkoff, 1999). However, the acquisition of the head-final or head-initial order as a result of the parametric setting in young infants remained to be addressed. As for the question of when a particular parameter is acquired, as mentioned earlier, the hypothesis proposed by Wexler (1998), Very Early Parameter setting, provides an answer. In VEPS, the order of the object and the verb is listed as one of the parameters that are set at very early stages. Advanced methodology and techniques make it possible to evaluate the parameter setting in young infants before the speech production stage. In this chapter an eye tracking experiment which tested 19-month old young infants on their knowledge of the head-final feature in their native language is reported. The chapter proceeds as follows: first, the literature on word order acquisition from different theoretical perspectives is reviewed. Then the original experiment in French on VO order acquisition, which the Hindi-Urdu experiment replicated, is reported. After the relevant background is provided, the experiment in Hindi-Urdu and its results are explained in detail and this is followed by the discussion.

### 4.2 WORD ORDER ACQUISITION THEORIES

Word order acquisition has been studied mainly from two different theoretical frameworks; one which claims that children are innately predis-

posed with the abstract syntactic properties that determine word order (Chomsky, 1981, 1995) and the other viewpoint which maintains that children employ lexical item based learning and thereby acquire word order (Tomasello, 1992). Vast amount of literature is available which argues for each of these theories with empirical evidence. A few relevant studies which support either one of these theories need to be reviewed in detail.

#### 4.2.1 *Word order acquisition by lexical learning*

The Verb-Island Hypothesis and the lexical learning theories (Tomasello, 1992; Akhtar, 1999; Tomasello, 2000; Dittmar et al., 2008) assume that the language acquisition is a product of an active learning process. According to this learned behavior approach, children do not have an innate abstract representation of syntax that sets parameters, but they learn the language in a piecemeal fashion from the environment. The Verb-Island hypothesis (Tomasello, 1992) proposes that it is around a verb that the early child grammar for word order develops. Children learn structures verb by verb with associated morphology and ordering constructions as templates for each particular verb. The word order acquisition according to the usage-based theory is nothing but an imitative learning of the ordering pattern of the arguments of a verb. The claim that the abstract properties of grammar are not innately present is put forth on the observation that during the early stages children do not extend the use of the learned verb constructions to different morphological forms or contexts than what they are exposed to through input. If children had access to the abstract rules, then they should be able to use a verb in all of its forms as soon as they learn one form of the verb. By observing that this is not the case, it is proposed that children learn the syntactic structures by an imitative learning method through rote memory. Children are said to be unaware of agent, patient, subject, object or any such abstract syntactic concepts, but are proposed to be operating on verb-specific concepts. For example, in English for the verb *kiss*, infants would learn that a *kisser* would precede the verb, and a *kissee* follows the verb, or for the verb *hit*, a *hitter* would precede the verb, and a *hittee* follows the

verb. This is the general model of learning assumed in the Verb-Island hypothesis. The hypothesis in Tomasello's own words:

"Until proved otherwise, we should assume that young children's early verbs and relational terms are individual islands of organization in an otherwise unorganized grammatical system. In the early stages the child learns about arguments and syntactic marking on a verb-by-verb basis, and ordering patterns and morphological markers learned for one verb do not immediately generalize to other verbs. The reason for this is that nascent language learners do not have any adult like syntactic categories or rules, nor do they have any kind of word class of verbs that would support generalizations across verbs. Processes of symbolic integration that serve to create sentences from words operate on a verb-specific basis as well. What children have at this stage are a knowledge of the roles played by various entities in these specific events, along with syntactic devices to indicate these."

(Tomasello 1992, p. 23)

Tomasello (2000) supports the Verb-Island hypothesis using observations gathered from comprehension and production studies. Tomasello analyzed spontaneous speech production of a native English speaking child, from 15 months to 24 months of age. The child was found to have used 162 verbs and their arguments during this time, but the construction types in which the verbs and arguments appeared were either in, as defined by Tomasello, type one, a complete SVO sentence, or type two, a phrase like V-PP-Obj. Inconsistencies were found in the use of the morphological markers, and in the use of the subject argument along with the verb. Verb production did not seem to be showing any continuity as an abstract categorical rule, that is, if one verb was produced in a novel construction, that did not result in another verb being used in that novel construction. No such generalization of application of word order knowledge was found. Rather verb constructions were found to be adhering to a lexical item-based development. Tomasello reports studies conducted in other languages as well to confirm the conclusions (Berman & Armon-Lotem (1995) for Hebrew; Serrat (1997) for



Catalan; Behrens (1998) for Dutch; Allen (1996) for Inuktitut; Gathercole et al. (1999) for Spanish; Stoll (1998) for Russian; Pizzuto & Caselli (1994) for Italian and Rubino & Pine (1998) for Brazilian Portuguese (from Tomasello (2000)).

The above-mentioned studies converged on concluding that children failed to produce novel utterances using learned verbs from their abstract syntactic categorical representation. The general claim is that the diversity, with respect to novel utterances in speech production, is limited to nominals in young infants. The children were found to be producing novel utterances by substituting concrete nouns in a sentence frame, but when it came to the argument structure of verbs children did not behave as if they had access to any syntactic structure rules.

According to Tomasello, there is no abstract representation of the syntax present in infants below 3-years of age. There is only a nominal slot available and the categorical representation of nominals. For verbs, the production was limited to structures which were already heard from the input. So it was claimed that if child grammar resembles adult grammar that is because children are learning language by imitating the adult production from their immediate environment and not because of the presence of an underlying innate system. Tomasello concludes from comprehension studies (Olguin & Tomasello, 1993; Akhtar & Tomasello, 1997; Tomasello et al., 1997; Dodson & Tomasello, 1998; Brooks & Tomasello, 1999; Tomasello & Brooks, 1999) that children under the age of 3 are incapable of reproducing novel verbs in novel utterances and that 3 years is the cut off age above which children start producing verbs in verb-general argument structures. The verb general knowledge appears in performance after 3-years of age as a result of children learning the verbs and their different constructions. The lexical-learning does not give any provision for children to have concepts of any abstract syntactic property that underlie grammar until age 3.

Moreover Tomasello reviews experimental studies conducted by various researchers (Maratsos et al., 1987; Pinker, 1987; Ingham, 1993) and claim that they did not report anything contrary to these proposals. The overarching assumption for these experiments was that when children are taught novel verbs, and if these verbs are taken up by an internal system, then children would immediately use these verbs in other struc-

tures with different syntactic constructions. The experiments found that children failed to produce novel verbs in new structures and that led to the conclusion that children only learn each verb and the linear order of its complements as a template. These studies were confirming the claim that children did not use verbs in different syntactic frames other than what they heard frequently in the input. In summary, the lexical-learning approach explains that children acquire verb general knowledge by an item-based learning.

Abbot Smith et al. (2001) examined the speech production of sixteen 2-year-old children and sixteen 3-year-old children to test the lexical-learning hypothesis. They used intransitive sentences with novel verbs and a known verb in two novel conditions and one familiar condition. In the novel conditions VS (83) and SV (82) orders were presented with the novel verbs, and in the familiar condition, a familiar verb in VS (84) order was presented. The presentation of the target sentences was embedded into a game activity.

(82) The cow *baffed*.

(83) *Meeked* the duck.

(84) Jumped the horse.

After repeatedly presenting the linguistic stimuli to children using puppets, the experimenter elicited production from children by asking questions like *What is happening?* The elicited productions were transcribed and analyzed later.

The results revealed that the two-year-old children corrected the VS order with the novel verbs to the canonical SV order only for 21% of the time, and the three-year-old children for 66% of the time. Comparing these results with Akhtar (1999), the authors argued that the performance based on word order knowledge increases with age. Contrary to their own claim, the 2-year-old children in their study corrected the VS order with a familiar verb to the canonical SV order for 72% of the time. This result is explained with the claim that it is a verb-specific knowledge and not an abstract knowledge about linear order. One of the main conclusions from this study is that the children below 2 years of age possess only lexical specific knowledge about the word order

properties of their native language. Contrary to what the item-based hypothesis could have predicted, i.e., 0% performance by the children in parsing non-canonical sentences with novel verbs, it was found that children performed correctly for 21% of the time. This discrepancy was not adequately addressed by Abbot et al.

Another point made by usage-based theories is that the frequency of the verbs in the primary linguistic data influences the learning of the verbs. The more frequent a verb in the input, the earlier the children learn the lexical items associated with that verb and their order. The argument is the same as mentioned before, that word order knowledge is an item-based specific knowledge and that there is no abstract knowledge of word order present in young children. So through frequent exposure to one verb, it becomes easier to memorize and learn. Matthews et al. (2007) conducted an experiment on two groups of forty-eight children with a mean age of 2.9 years and 3.9 years. They used four high-frequency verbs (*push, pull, throw, wipe*), four medium-frequency verbs (*shove, flip, drag, rub*) and four low-frequency verbs (*ram, tug, hurl, dab*). The children were taught to produce these verbs in a training phase in non-canonical SOV order. The subjects were expected to answer using the canonical word order when a question like *What happened there?* was asked. The results revealed that there was an effect of verb-frequency on the performance of young children in correcting to non-canonical order to the canonical order. The two-year-old subjects corrected the SOV order to SVO order more in the sentences with the high-frequency verbs than with the low-frequency verbs ( $p \leq 0.001$ ). No effect of verb-frequency was found in the performance of the three-year-old children. As the younger children could perform better only in conditions with the high-frequency verbs, this led to the conclusion that they learn word order in a gradual manner through individual lexical items.

The general conclusion from the studies is that children learn language by imitating the adult input in young ages. As the children grow old, they accumulate enough verb-specific knowledge to generate a verb general rule for word order. The language learners at early ages are said to be dependent on lexically specific syntagmatic and paradigmatic categories. According to this theory, the word order knowledge which children display at young ages is not based on the abstract knowledge of

word order of the native language but merely the reproduction of memorized verb-specific order of arguments. This type of imitative learning is said to continue up to the third year of life. According to this approach, children are taxed with the task of learning the language rather than by acquiring language as a biological process. The theory is based on the assumption that children need to learn the lexicons in order to acquire the word order rules pertaining to each language. The Principles and parameters approach, on the other hand, places its axis on the presence of an innate system of the human mind or brain which aids in language acquisition as an automatic process.

#### 4.2.2 *Word order acquisition guided by innate knowledge*

The Principles and Parameters approach takes into account the underlying universal similarities of grammar and explains language acquisition from this dimension. The framework of UG proposed that children are born with a language acquisition device that is grammar sensitive, which detects structures and language specific properties from the input and matches it with the innate rules in the system and thereby sets various parametric values. There are different accounts on how these structures are identified by the acquirers to set the parameters. Answering the question of how parameters are set is important if one wants to endorse the P&P approach, as this is one of the main criticisms the P&P approach face. When the innate system is stimulated with an input, there must be cues within the input that facilitate the computational system to parse the input and set the parameters accordingly. Hence one of the tasks linguists face is to identify these cues and evaluate them cross-linguistically. According to Chomsky (1993), language is acquired by:

“a process of selection of a rule system of an appropriate sort on the basis of direct evidence. Experience yields an inventory of rules through the language acquisition device of the language faculty.” (Chomsky 1993, p. 641)

So what constitutes direct evidence? For the acquisition of word order, what are the cues that children get from their direct input, which is the

link between the input and innate abstract rules (Pinker, 1994; Mazuka, 1996)? Before going into the various cues in the input that aid in parameter setting, first a few accounts on the role played by word order in learning verb meaning are explained.

The syntactic bootstrapping hypothesis (Landau & Gleitman, 1985; Naigles, 1990; Fisher et al., 1994; Fisher, 2002; Fisher et al., 2010) is one of the proposals that tries to explain how children acquire the meaning of verbs through word order, structure of arguments, and predicates. The hypothesis states that children make use of the syntactic frame to decipher the predicate and its meaning. Young learners use the syntactic frames of sentences initially to map a word to its meaning and thereby start sentence comprehension by processing the syntactic frame of simple verb constructions. As children hear an utterance, they collect information about the syntactic structure of predicates, and about the number of arguments it permits. The nouns are mapped one-to-one in an event by doing an analysis of the syntactic frames. Children are said to be born with this mental knowledge.

Naigles (1990) argues that the sentence structure is a powerful source to learn verb meaning. A study was conducted on twenty-four children (mean age 2;1 years) to test this hypothesis. The children were taught novel verbs in transitive (85a) and intransitive frames (85b) separately.

(85) (a) Look! The duck is *gorping* the bunny.

(b) Look! The duck and the bunny are *gorping*.

After a training phase it was tested whether the children who heard the novel verb in a transitive frame correlated the verb to a causative action and the children who heard the novel verb in an intransitive frame correlated the novel verb to a reflexive action. A significant effect in the looking time towards the matching screen ( $p = 0.001$ ) for the corresponding sentences was found from the results. This was provided as empirical evidence for the hypothesis that children possess the knowledge of syntactic frames and they make use of this knowledge to learn verb meanings.

Fisher (2002) studied the comprehension of transitive and intransitive sentences using pseudo-verbs and pronouns in children. The use of pronouns ensured that children had the structural cues alone about the

verb and that the children did not pick up the agent and patient information from the nominals. Across two experiments 48 children were tested. The subjects were divided into a younger group with a mean age of 28.6 months and an older group with a mean age of 31.1 months for the analysis. Videos of causative action were shown to children, which consisted of target sentences embedded in a script. During the video presentation, the experimenter asked questions and made children point to the agent of the action. Children were exposed to the pseudo-verbs in either a transitive sentence structure or an intransitive sentence structure. The sentences presented were the following:

- (86) (a) She *stipes* (her) over there.
- (b) She *braffs* (her) over there.
- (c) She *pilks* (her) back and forth.
- (d) She *gishes* (her) around.

It was found that the children who heard the transitive sentences chose agents correctly more often than the children who heard the intransitive sentences. Since the sentences had no nominals but only pronouns it became evident that the children chose the agent as the one who is causing the event using the linear order of the sentences alone. The children identified the number of arguments, and when there were two nouns they considered the predicate to be transitive. This gave evidence that children as young as 28 months are influenced by the abstract syntactic properties of grammar.

Along similar lines, one of the experiments by [Gertner et al. \(2006\)](#) was on testing comprehension of transitive sentences with novel verbs by 25-month old children. Subjects were divided into two groups; one group received sentences with a pair of characters, duck and bunny, as the agent and the patient (87), and the other half received the same characters in reversed roles (88) (test sentences in Naigles, 1990 were of transitive and intransitive frames with the same pseudo-verb).

(87) The duck is *gorping* the bunny.

(88) The bunny is *gorping* the duck.

All the characters were introduced in a familiarization phase. After familiarization, in a training phase, the subjects were shown videos where the characters carried out actions representing familiar verbs, hug and feed, along with sentences like *Look, the duck is hugging the bunny*. In the video, two clips were shown: one in which the action matched the sentence and another clip where the same characters were involved in a different action. The training phase was followed by the test phase. Linguistic stimuli were given with pseudo-verbs which corresponded to different actions; one where a character wheeled another character in a wagon, and another where one character tipped the other character in a rocking chair.

The experiment was video recorded, and the visual fixation duration was calculated. Children were found to have inferred that *gorping* is something that the agent is doing to the patient from the word order of the sentence. The same methodology, with which the 25-month old children were tested, was used to test the comprehension of word order in 21-month old children, and similar results were found. Here it has to be noted that the subjects were given a training phase, and one of the criticisms (Dittmar et al., 2008) that rose was that the training phase prepared children to perform adequately during the test phase. This criticism is discussed later with the results from the Hindi-Urdu experiment.

A study by Lidz et al. (2003) provides evidence for the innateness of grammar using constructions which are scarce in primary linguistic data. (Child-directed speech was analyzed by the authors, for complete results see the original paper.) They tested the comprehension of the anaphoric construction using *one*, in 18-month old English speaking children. During the training phase a single object was presented, and during the test phase, another similar object but in a different color was presented along with the previously presented object. A yellow bottle was presented first and later two bottles, one yellow and one blue, were presented and the children were asked *Now, do you see the other one?* The duration of the looking time was calculated using a frame by frame analysis. The results showed that the children interpreted the anaphoric structure correctly. This furnished a strong argument against the imitative learning argument. The primary linguistic data is impoverished



for the presence of the anaphoric construction. Yet very young children have been found to have the knowledge about this construction, both for comprehension and for assigning new meanings to novel verbs. Along the lines of results provided by Lidz et al., many other studies provided evidence for the presence of syntactic structures which are scarce in the primary linguistic input (Göksun et al., 2008; Lee & Naigles, 2008; Matsuo et al., 2012). Evidence for the comprehension of structures which are infrequent in the linguistic input disproves the core claim by the lexical-learning hypothesis.

Syntactic bootstrapping requires a child to have some amount of linguistic knowledge in terms of locating the categories in the speech signal. Children should identify the nouns/arguments in a syntactic frame to choose the word order. The knowledge of a transitive verb frame is needed to construct the word order for a transitive verb and to assign causality to it. So how can a child without any formal training and adequate maturation of cognition, isolate nouns and verbs in the incoming speech signal?

This question brings the discussion back to the cues present within the speech input that are used for the setting of parameters. The speech signal which young infants are exposed to contains various acoustic cues. Before any semantic, syntactic or pragmatic skills emerge, infants are found to be sensitive to the phonological and the prosodic cues within the speech signals. It has been found that the newborn brain is sensitive enough to differentiate linguistic stimuli from non-linguistic stimuli (Dehaene-Lambertz et al., 2002). The prosodic bootstrapping hypothesis was put forth by Morgan et al. (1987). The assumption was that there are cues in the speech input that facilitate the identification and grouping of words into phrases. Once these cues are picked up and analyzed by the internal grammar, they allow the setting of parameters based on abstract principles. The benefit of this proposal lies in the fact that it does not tax the infants with a necessity to have prior knowledge of lexical items, their meanings, and syntactic relations.

Mazuka (1996) identifies a paradox within the parameter acquisition approach: for a child to set the parameter which determines the linear order of head and complements, the child should be able to identify which are heads and which are complements in the input. The paradox



was earlier termed the linking problem (Pinker, 1994). In a natural linguistic environment, the identification of heads and complements is difficult for the newborn infants because strings of words do not contain a tag as head or complement with which the child can identify the heads and complements. Mazuka argues for two basic parameters that determine the basic configuration of the word order in any language, which are: the branching direction parameter and the head-directionality parameter. For the branching direction parameter children do not need to analyze the small segments in the input data. Instead, they can rely on the prosodic cues that distinguish phrases and clauses. So if a child can identify the clause boundaries he or she can isolate the rule of branching direction and set the parameter accordingly. The branching direction parameter is defined as the rule that sets the direction of recursively embedded clauses (Lust & Chien, 1984). So languages like English and French are right branching languages (89a) and languages such as Japanese and Hindi-Urdu are left branching languages (89b). The head-directionality parameter and the branching direction parameter differ in their values in languages like German and Chinese. The right branching German is head-final and the left branching Chinese is head-initial.

(89) (a) [I met the teacher [who called the student [who had an accident ]]].

(b) [[[Ziko-ni atta] gakusee-ni denwasita] sensee-ni  
 had an accident student-DAT called  
 atta].  
 teacher-DAT met  
 'I met the teacher who called the student who had an  
 accident.'

(Examples taken from Mazuka, 1996)

The prosodic cues about clause boundaries are proven to be one of the most prominent cues in infant-directed speech (Fisher & Tokura, 1996). The literature gives evidence that infants as young as 4.5 months are sensitive to the prosodic clause boundaries, and Mazuka claims that this would be one of the earliest cues to syntactic processing that children

use. Once the children set the branching direction parameter using the prosodic clause boundary cues, subsequently the position of head and complements can also be set. One of the main underlying assumptions of the prosodic bootstrapping hypothesis is that children use prosody to initiate language acquisition process.

However, contrary to Mazuka's claim, evidence has been obtained that even smaller units of speech, like words, are prosodically distinguishable. The claims that the young infants are sensitive to phrase boundaries, word boundaries, and intonation boundaries are validated with empirical evidence (Christophe et al., 1997; Guasti et al., 2001; Christophe et al., 2003, 2008; Gervain et al., 2008).

Morgan & Demuth (1996) introduced the term phonological bootstrapping to the proposal of prosodic bootstrapping, as the primary linguistic data contains more information like phonetic, phonotactic, prosodic and stochastic cues than just prosodic cues.

The main criticism the prosodic bootstrapping hypothesis received was that there is no one-to-one correlation between syntax and prosody. So children may not be able to set the parameters solely based on prosodic or phonological cues, yet it is said to provide with a start. Even for the languages with free word order, children as early as 3-years show the knowledge of the basic unmarked order, SOV, in comprehension and in production (Otsu, 1994; Sugisaki, 2008). This indicates that there should be means through which children can identify nouns and verbs from the input from early on. Only after identifying these categories can there be a relation built on their order. When there are nouns that precede verbs in a considerable amount of time in the primary linguistic data, and if the children can identify the nouns and verbs as different categories, then that should enable the language acquisition device to set the parameters for the OV/VO order accordingly.

Refraining from adopting any bootstrapping hypotheses, there are studies in the literature which assumed the UG framework in general and investigated the acquisition of word order and found that it is acquired early. In pioneering work, Hirsh-Pasek & Golinkoff (1999) investigated whether the young infants in one-word production stage comprehend the grammatical significance of word order in their native language, by employing the preferential looking paradigm. Forty-eight

infants within the age range of 16 to 19 months (mean age 17.5 months) were tested. The subjects were shown video clips with auditory stimuli. Character familiarization was carried out in a training phase. During the experimental phase reversible transitive sentences (90) were presented while two video clips appeared on the screen: one in which the agent and patient matched the test sentence and the other in which agent and patient roles were reversed. Matching and reversed clips were counter-balanced. Four verbs were used: *hug*, *wash*, *tickle*, *feed*.

(90) Big Bird is washing Cookie Monster.

Visual fixation duration was calculated and analyzed. The significant result from this study is that the infants preferred looking at the clip that matched the linguistic stimuli. That led to the conclusion that infants at 17 months of age are sensitive to word order. The children correctly identified the agent and patient from the word order. The results give clarity towards the fact that the children are aware of syntax and word order from early stages of life. The argument against this result from researchers who favor the usage-based theory was that it was the lexical knowledge of the verb that enabled the children to perform accurately.

Sugisaki (2008) analyzed corpus data from young Japanese children and found that by the age of 2.5 years Japanese children have acquired the OV order and use it in production. SOV is the basic word order in Japanese and SVO is a marked order with restrictions. In embedded clauses, in idiom chunks and in direct object *wh* questions SVO is ungrammatical. The prediction was that if children had the knowledge of markedness and that Japanese is a head-final language then the children would produce these structures abiding the word order restrictions, in the manner similar to adult grammar. If children accepted both VO and OV orders in the absence of parameter setting, then they would not display syntactic constraints in VO order. Sugisaki focused on the direct object *wh* questions (91 a) to investigate this particular observation.

- (91) (a) *Eri-ga nani-o tabeta (no)?*  
           Eri-NOM what-ACC ate Q  
           ‘What did Eri eat?’  
       (b) \**Eri-ga tabeta (no), nani-o?*  
           Eri-Nom ate Q what-Acc

'What did Eri eat?'

The spontaneous production corpus of four children was analyzed, and it was found that the children used OV order in direct object *wh* questions. From the 436 sentences analyzed there was only one instance of violation of the OV order which led to the conclusion that children are aware of what is the basic word order configuration and the various structural constraints associated with it.

Isobe (2009) studied the acquisition of Head Internal Relative Clauses (HIRC) and Head External Relative Clauses (HERC) in Japanese children. The children's knowledge to differentiate between the HIRC and HERC, and the role of prosody in disentangling ambiguity were addressed in the study. The subjects tested were 17 Japanese children within the age range of 3;3 - 5;4 years. The subjects were divided into two groups; Group A and Group B. Both groups received animated stories and at the end of the story children had to perform a truth value judgment task about a sentence which a puppet produced. Group A and B received one common type of sentence; a structurally ambiguous sentence with HIRC prosody, which could either be interpreted as HIRC using the prosody (93 a), or as HERC as that is a syntactically possible interpretation (93 b).

- (92) Sample Story 1: A panda, a rabbit, and a monkey all went to play outside. But after a period of time, it suddenly started raining. The rabbit said, "I have an umbrella" and opened it. The panda got under the rabbit's umbrella, but the monkey couldn't. When they all returned home, the monkey was wet all over. The panda brought a big towel for the monkey, and dried him.

There were two possible interpretations for the test sentence:

- (93) (a) *pro [pandas-an-ga nurechatta-no]-o kawakashi-ta*  
           *pro the-panda-Nom got-wet-Comp-Acc dry-made-Past*  
           *yo.*  
           Excl  
           '(Someone) made dry the panda, who got wet.'
- (b) *Pandas-an-ga [nurechatta-no]-o kawakashi-ta yo.*  
       *the-panda-Nom got-wet-one-Acc dry-made-Past Excl*  
       'The panda made dry the one who got wet.'

In this task, if the children identified (93b) as false in the judgment, then that would indicate that children interpreted the sentence as HIRC.

In another task, Group A and B received two different types of sentences. Group A received a structurally unambiguous sentence with HERC prosody which could only be interpreted as HERC, and an ambiguous sentence with HERC prosody which can have two different interpretations. The difference in the interpretation is caused by the topic marker *wa* in the sentence that unambiguously marks a matrix clause subject. So the interpretation of the sentence presented to Group A is given in example (95). But the sentence Group B received could have two interpretations, an expected HERC interpretation (96) and a syntactically possible HIRC interpretation (97).

(94) Sample Story 2: A monkey, a panda, and a rabbit are playing outside. Suddenly it got dark and started raining. The rabbit said, "I have an umbrella" and opened it. The monkey got under the rabbit's umbrella, but the panda wasn't able to. When they all returned home, the panda was wet all over. The monkey brought a hair dryer for the panda, and dried him.

(95) *Pandas-an-wa [nurechatta-no]-o kawakashi-ta yo.*  
the-panda-Top got-wet-one-Acc dry-made-Past Excl  
'The panda made dry the one who got wet.'

(96) *Pandas-an-ga [nurechatta-no]-o kawakashi-ta yo.*  
the-panda-Nom got-wet-Comp-Acc dry-made-Past Excl  
'Panda made dry the one who got wet.'

(97) *pro [pandas-an-ga nurechatta-no]-o kawakashi-ta yo.*  
pro the-panda-Nom got-wet-Comp-Acc dry-made-Past Excl  
'(Someone) made dry the panda, who got wet.'

The results indicated that both groups performed above chance for ambiguous sentences with HIRC prosody, as expected. Group A performed above chance for structurally unambiguous HERC sentence. However, the Group B children, on average, processed ambiguous HERC prosody sentences as HIRC sentences which suggested that children did not rely on prosodic cues alone. The authors conclude that the Japanese children

by the age of 3 have the knowledge of HIRC and HERC constructions and that while faced with ambiguous sentences they do not rely on prosodic cues alone.

Combining different methodologies like the intermodal preferential looking method and the weird word order paradigm, used in studies which favored innateness and which favored lexicalized learning respectively, [Franck et al. \(2011\)](#) investigated the knowledge of the abstract word order in young children. The studies which had previously employed the preferential looking method claimed that abstract syntactic rules are innately present. This claim was made by virtue of the fact that young children preferred to look at actions that corresponded to the linguistic stimuli. On the contrary, studies that had employed the weird word order (WWO) claimed that children learned the word order in an item-based manner as they failed to correct ungrammatical word orders with pseudo-verbs to the target language word order. Franck et al refuted this claim through empirical evidence which was gathered from a study that was designed with both of the above-mentioned methods (the experiment replicated and extended [Matthews et al. \(2007\)](#)).

Twenty-four native French speaking children participated in the study. Children were divided into two groups; 12 children in the age range of 2;04 - 3;03 years (mean age - 2;11) and 12 children in the age range of 3;04 - 4;04 years (mean age - 3;11). Four verbs (*pousser, taper, laver, mordre*) and four pseudo verbs (*pouner, touser, nuver, daser*) were used and 64 sentences were produced out of the 8 verbs. The pseudo-verbs corresponded to actions which were not lexicalized in French: bumping into someone's belly with the head, catching someone with a strainer, pulling someone's tail, putting a crown on someone's head. Two verbs and two pseudo-verbs were presented in grammatical NP-V-NP order ((98), (99)) and in the ungrammatical weird word order respectively ((100), (101)). V-NP-NP and NP-NP-V were the weird word orders employed.

(98) *La vache mord le chien.*  
the cow bite the dog  
'The cow bites the dog.'

(99) *Le lion dase le cheval.*  
the lion pseudo-verb the horse

'The lion pseudo-verb the horse.'

- (100) *La vache le cheval lave.*  
 the cow the horse wash  
 'The cow the horse washes.'

- (101) *Le chien la vache nuve.*  
 the dog the cow pseudo-verb  
 'The dog the cow pseudo-verb.'

Two experiment lists were made with each weird word order. The stimuli were presented to the subjects in two parts, in the first part three videos with linguistic stimuli were presented, and a fourth video was presented without any linguistic stimuli. Children were asked to explain what was happening in the fourth video, to a puppet which had its eyes covered. In the second part, one video with the linguistic stimulus and one video without the linguistic stimuli for eliciting verbal production were presented. If children produced the exact sentence structure with the same verb, it was coded as a match. If children corrected the weird word order to the canonical order along with the same verb, it was coded as a reversion, and corrections with a different verb were coded as other sentences.

The significant results from this study are: the older children and the younger children performed identically in correcting the ungrammatical weird word order into canonical SVO. The effect of grammaticality (SVO/weird word order) and lexicality (verb/pseudo-verb) were not different in young children and older children. Since there was no significant difference in the performance between the two age groups, the results favored the hypothesis that abstract word order knowledge is present in young children.

Evidence from word order acquisition studies is conclusive on the fact that children exhibit the knowledge of basic word order in both expression and comprehension as early as 21 months of age. However, all the criticisms which were raised, by researchers who support the lexicalized learning theory, against the studies that claimed the innate knowledge of abstract syntax were not addressed in any of the above-mentioned literature. A detailed description of these criticisms, and about a study

that addressed these issues, which served as the background for the Hindi-Urdu acquisition experiment, are given in the next section.

#### 4.3 THE ACQUISITION OF VO ORDER

Franck et al. (2013) designed a methodology which combined the use of pseudo-verbs, the weird word order paradigm, the preferential looking, and a procedure devoid of any training phase for the test items. The pseudo-verbs in the test sentences ensured that the children were not performing based on the prior knowledge of the lexicon or semantics of the verb. The weird word order paradigm (Akhtar & Tomasello, 1997) checked the knowledge of the grammaticality of word order by virtue of contrasting a canonical order with an ungrammatical order. There was no training phase involved in the study and that eliminated the chance for children to learn the novel verb or the role of agent and patient presented in test sentences. Thus the subjects in the study did not perform based on a learned template. These were the main criticisms that were raised against the studies that claimed innate abstract knowledge of word order in infants (criticisms raised in Dittmar et al. (2008)). Some other variables were controlled in the Franck et al.'s methodology: the grammaticality was controlled using transitive sentences, instead of transitive and intransitive comparisons as in Naigles (1990). Causative and non-causative actions of the same activity were used in the video clips as target and distraction as opposed to the agent-patient role reversed causative action in Fisher (2002); Gertner et al. (2006).

Nineteen 19-month old French acquiring infants served as subjects for the experiment. In accordance with the phonotactic and phonological features of French, two bisyllabic pseudo-verbs: *daser* and *pouner*, were created. These verbs corresponded to non-lexicalized actions in French, *pouner* illustrated an agent catching a patient's head under a net, and *daser* illustrated an agent putting a crown on a patient's head. *Pouner* was used in the grammatical NP-V-NP order (102 a), and *daser* was used in the ungrammatical NP-NP-V order (102 b). The pseudo-verbs were 80% and 66.7% phonologically representative of the transitive verbs in French according to the statistics computed from the French database

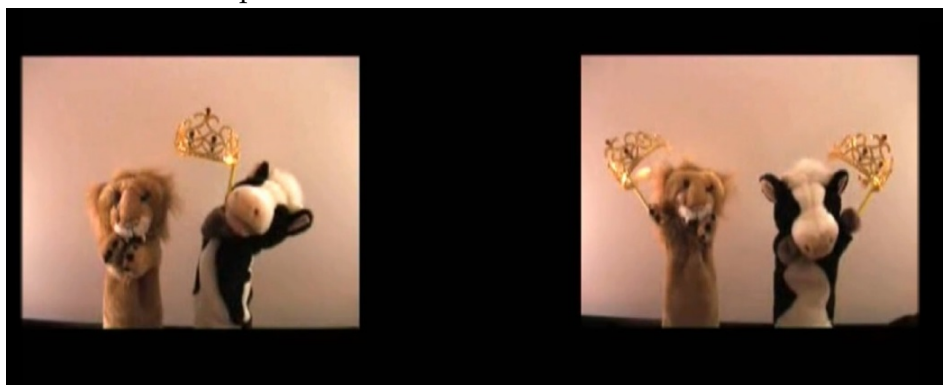


Lexique. A total of six sentences were used, 3 in grammatical (NP-V-NP) and 3 in ungrammatical (NP-NP-V) order.

- (102) (a) *Le lion poune le cheval.*  
           the lion pseudo-verb the horse  
           'The lion pseudo-verb the horse.'
- (b) *La vache le lion dase.*  
           the cow the lion pseudo-verb  
           'The cow the lion pseudo-verb.'

Animal puppets were used as characters in the video clips. The characters, and different positions on screen where clips would appear and the children were familiarized to the novel actions in the introductory phase. Neither the pseudo-verbs nor any transitive sentences were provided during this phase. The infants were asked questions like *What is happening?*, *Do you see it?*, *What is it?* etc. Each sentence was presented in a video with two clips (see Figure 13). In one clip an agent performed a causative action on the patient, and in another clip both characters performed the same action on themselves synchronously. The positions of causative and reflexive actions were counterbalanced. Each video was divided into five 4-second frames. The first frame consisted of a base-line sentence, used for grabbing the subject's attention, like *Look, what is happening?* The linguistic stimuli were presented three times in a loop for the next 16 seconds of the video.

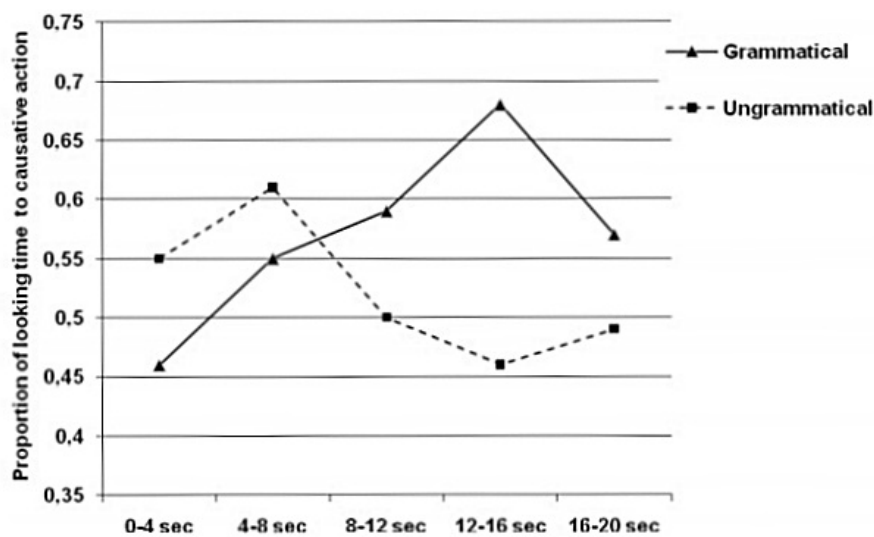
Figure 13: Video clips depicting the causative and non causative action in the French Experiment



The predictions were that, if children processed word order by virtue of abstract syntactic knowledge then there would be a preference to the causative action upon hearing the grammatical NP-V-NP order and no preference would be shown during the ungrammatical NP-NP-V order. If children learned word order through lexicalized learning then no preference would be shown for the causative or the non-causative action, neither in the grammatical nor in the ungrammatical condition.

The results showed above chance looking time to the causative action for the grammatical condition,  $p < .01$  during the 12-16s frame of the video. Children preferred the causative action over the reflexive action when the grammatical word order was presented. During the ungrammatical condition there was no preference observed and the results were at chance level (see Figure 14).

Figure 14: Proportion of mean looking time during the grammatical and ungrammatical condition (Franck et al., 2013)



Franck et al.'s experiment successfully proved that at the pre-verbal stages children have the knowledge of the VO order. Since there was a significant preference in looking at the causative action upon hearing

the grammatical sentence and no preference while hearing the ungrammatical sentence, the prediction that the children possess innate abstract knowledge of syntax was borne out. Children did not have the semantics of the verb to deduce agent and patient as *pouner* and *pounee*. If the children interpreted the first NP as the agent, then there would have been an equal preference for the grammatical and the ungrammatical orders, as both the grammatical and the ungrammatical word orders used the agent-patient order. Thus results point to the fact that children did not simply interpret the first NP as the agent, and that they parsed the whole sentence using the knowledge of word order alone.

In order to cross-test the results from the French study in a language with another word order, the same methodology was employed in Japanese. Omaki et al. (2012) tested 48 infants in three age groups: 24 infants within the age range of 18.2-20.6 months (mean=19.0 months), 24 infants within the age range of 27.0-28.4 months (mean=27.5 months) and 25 infants within the age range of 31.7-33.0 months (mean=32.3 months). Two pseudo-verbs were used, *neketteru* and *isetteru* in the grammatical NP-NP-V condition (103 a) and in the ungrammatical NP-V-NP condition (103 b).

- (103) (a) *Wanchan-ga nekochan-o neket-teru.*  
           dog-NOM   cat-ACC   pseudo-verb  
           'The dog the cat pseudo-verb.'
- (b) *Ushisan-ga iset-teru           osarusan-o.*  
           cow-NOM   pseudo-verb   monkey-ACC  
           'The cow pseudo-verb the monkey.'

The same methodology as in Franck et al. was used. However, different results were obtained for the Japanese study than the French study. The 19-month old infants showed a preference for the reflexive action for the NP-NP-V sequence, and no preference was found for the NP-V-NP sequence. The 27-month old infants showed no preference to any of the actions in both word orders. The 32-month old infants showed a clear preference for the causative action during the grammatical NP-NP-V sequence. There was no clear interpretation of why only the 32-month old infants parsed grammatical sequences. It was suggested that since Japanese is a pro-drop language, the children at 19 months do not have

sufficient NP-NP-V sequences in their primary linguistic data to set the parameter.

In order to further investigate the results from Franck et al. and Omaki et al., the same methodology was applied with slight modifications in another predominantly head-final language, Hindi-Urdu. Unlike French but like Japanese, Hindi-Urdu is a language which allows freedom of word order. So the primary linguistic input children receive consists of not just the OV order but different orders like VSO, VOS, OVS, SVO. Even though the unmarked order is head-final, children are faced with sentences with mixed word orders. It would be particularly insightful to analyze how children, who acquire a language which has freedom of word order, and null arguments, perform in this test and how far ahead is their syntactic knowledge at the age of 19 months.

#### 4.4 THE ACQUISITION EXPERIMENT OF OV ORDER IN HINDI-URDU

##### 4.4.1 *Rationale*

Taking into account Very Early Parameter Setting hypothesis, the main rationale behind this experiment was formulated, which was that the parameters determining the OV order are set in children who are exposed to a head-final language, Hindi-Urdu, at 19 months of age. To test this, Franck et al.'s experiment was replicated with slight modifications (which are detailed in subsequent sections). The predictions for this investigation were that, based on the parameter setting for the OV order, the children would be able to parse the grammatical word order. Children would not exhibit any preference for causal action or reflexive action when the ungrammatical weird word order is presented.

##### 4.4.2 *Experimental Design*

###### 4.4.2.1 *Participants*

Twenty infants within the age range of 1;7,2-1;7,7 years (mean age: 1;7,4 years) participated in the study. All the infants were exposed to Hindi-

Urdu from birth as their mother tongue. All the subjects were native Pakistanis who were born and raised in Barcelona city. Due to cultural reasons they had little exposure to the local languages. There was no family history of any language, cognitive or hearing impairment. No child was excluded from the study.

#### 4.4.2.2 Test sentences

Two pseudo-verbs were created, after collecting the commonly occurring phoneme combinations, from a list of most common Hindi-Urdu verbs: *chon* and *khala*. The pseudo-verb *chon* ended in a consonant and its simple past tense was conjugated as *chona* by adding the perfective participle *-a* to the verb root. The pseudo-verb *khala* ended in a vowel and its simple past tense was conjugated as *khalayaa*, by adding *-y* and *-a*. The subjects of these verbs were given the ergative case *-ne* and all the objects, which were animate, were given the accusative marker *-ko*. The pseudo-verb used in the grammatical condition was the trisyllabic *khalayaa*, and for the ungrammatical condition the bisyllabic *chona* was used.

The word order sequences were not replicated exactly from the Franck et al.'s methodology. As it was shown from the word order frequency analysis of Hindi-Urdu spontaneous speech, V-NP-NP was the least occurring non-canonical word order, and so that was chosen for the ungrammatical condition. NP-NP-V was the grammatical sequence and V-NP-NP was the ungrammatical sequence. Test items contained a total of six sentences, 3 grammatical (104 a - 104 c) and 3 ungrammatical (104 d - 104 f).

- (104) (a) *gay-ne bakri-ko khalayaa.*  
           cow-ERG sheep-ACC pseudo-verb  
           'The cow the sheep pseudo-verb.'
- (b) *kuthe-ne gadhe-ko khalayaa.*  
           dog-ERG donkey-ACC pseudoverb  
           'The dog the donkey pseudo-verb.'
- (c) *sher-ne ghode-ko khalayaa.*  
           lion-ERG horse-ACC pseudoverb  
           'The lion the horse pseudo-verb.'

- (d) *chona*            *gay-ne*        *sher-ko*.  
      pseudo-verb cow-ERG lion-ACC  
      'pseudo-verb the cow the lion.'
- (e) *chona*            *gadhe-ne*        *kuthe-ko*.  
      pseudo-verb donkey-ERG dog-ACC  
      'pseudo-verb the donkey the cow.'
- (f) *chona*            *bakri-ne*        *ghode-ko*.  
      pseudo-verb sheep-ERG horse-ACC  
      'pseudo-verb the sheep the horse.'

The sentences were spoken by a female native Hindi-Urdu speaker from Peshwar, Pakistan and was recorded at a sound proof lab in the Universitat Autònoma de Barcelona. The speaker was given instructions to use the same prosody in all the sentences. A practice session was given to the speaker. All the sentences were recorded in neutral focused declarative intonation contour of Hindi-Urdu (105).

(105) [[L\*H] [L\*H] [H\*L]]

This rendered the V-NP-NP sequence ill-formed.

### *Materials*

The entire experiment consisted of training videos and experiment videos. The training videos were used to introduce the animal puppets (horse, donkey, lion, sheep, dog and cow) and for familiarization. The experiment videos consisted of six videos of target clips (with target frames and distractor frames), a blank screen and a cartoon clip in between the six target clips. The linguistic stimuli were presented through speakers. The distracting cartoon clips were also in Hindi-Urdu. The story of *Chulbuli* (see Figure 15), a familiar cartoon for native Hindi-Urdu children, was segmented and added before and after each experimental video.

Figure 15: Chulbuli cartoon clip



#### 4.4.2.3 Instrumentation

Testing was carried out using a Tobii 120 eye tracker, version 2.0. The screen size was 17 inches and the distance from screen to eye was 65 cm. Infant control calibration was used. All the tests were conducted in the Psycholinguistics lab at the Facultat de Psicologia, Universitat de Barcelona.

#### 4.4.2.4 Procedure

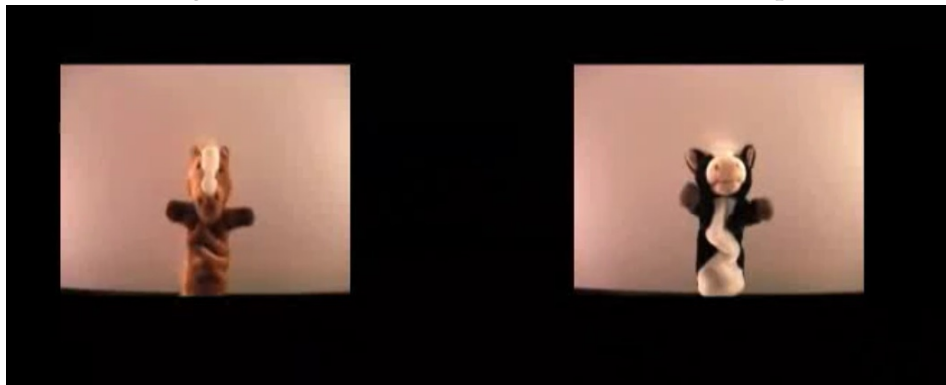
Each child was made to sit on the lap of the mother/care taker. The mother/care taker had her eyes closed during the experiment phase but not during the familiarization phase. Two sets of videos were presented during the familiarization phase. In the first set each puppet appeared on either side of the screen accompanied by an auditory stimulus, like, *Look! a cow, hello cow!* (see Figure 16).

Figure 16: The character introduction video clip



In the second set two puppets appeared on the left and right side of the screen, accompanied by an auditory stimuli targeted towards one character, *Look where is the cow, do you see it?* (see Figure 17).

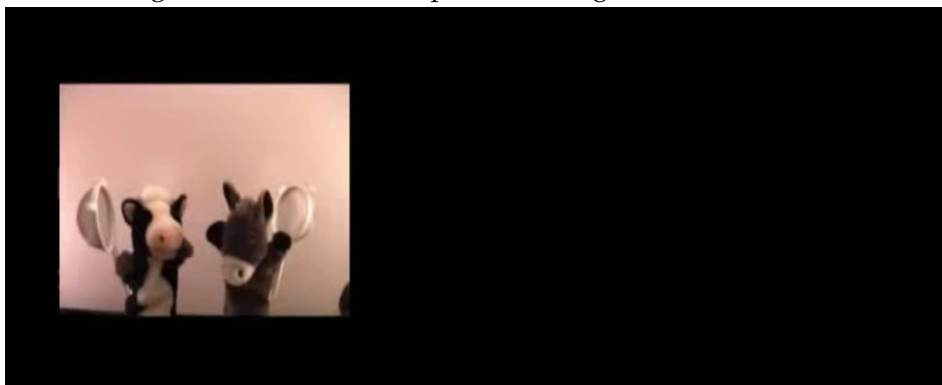
Figure 17: The character identification video clip



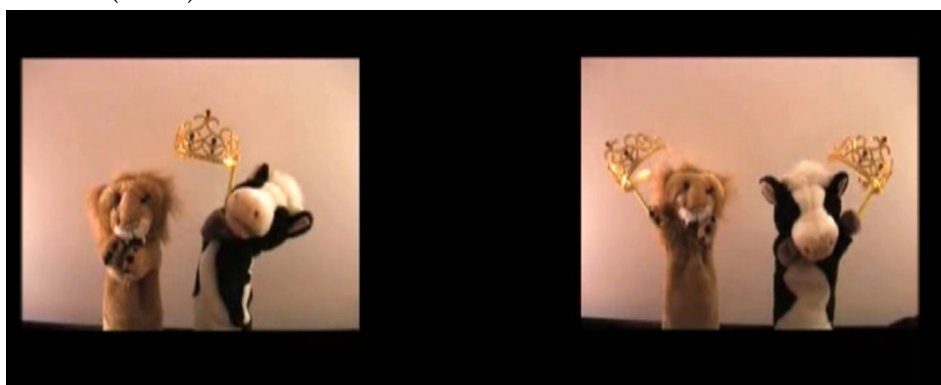
Then the novel action was introduced with one clip on either side of the screen with a question, *what is happening* or *Do you see that?* (see Figure 18), as auditory stimuli.



Figure 18: The video clip introducing the novel action



No transitive sentences with the two character names or any sentences with the pseudo-verbs were given to children during the familiarization phase. A short break was given before the experimental phase. Infant calibration was carried out prior to running the tests after the break. The experimental phase contained six videos. In each video, two clips in which the same characters involved in a causative and reflexive mode of the same action appeared on the right and left side of the screen. The position of causative and reflexive actions were counter balanced. The grammatical and the ungrammatical sentences were presented in random order. The ungrammatical V-NP-NP order sentences were presented with the pseudo-verb *chona* as in Figure 19.

Figure 19: An example from the ungrammatical sentence video sequence (*chona*)

Grammatical NP-NP-V order sentences were presented with the pseudo-verb *khalaya* which depicted catching a character's head under a net (see Figure 20).

Figure 20: An example from the grammatical sentence video sequence (*khalayaa*)



The presentation of the target clips was controlled by the experimenter. When children were not looking at the screen or distracted, the experimenter paused the video and tried to get the attention back on the screen by asking questions like *Oh where are the animals?*, or *What is happening?* or *Something just happened on the screen!* etc. During the presentation of the cartoon clips the experimenter prompted the child to remain looking at the screen by making relevant remarks about the events in the cartoon. However, neither the experimenter nor the care taker interacted with the subjects in any mode during the test clips. The children were given a soft-toy by the end of the experiment as a gift for being attentive. The soft-toy was used as a reinforcement. Children were told that if they sat and watched the entire video they would receive this toy by the end. Preferential looking was the expected response mode.

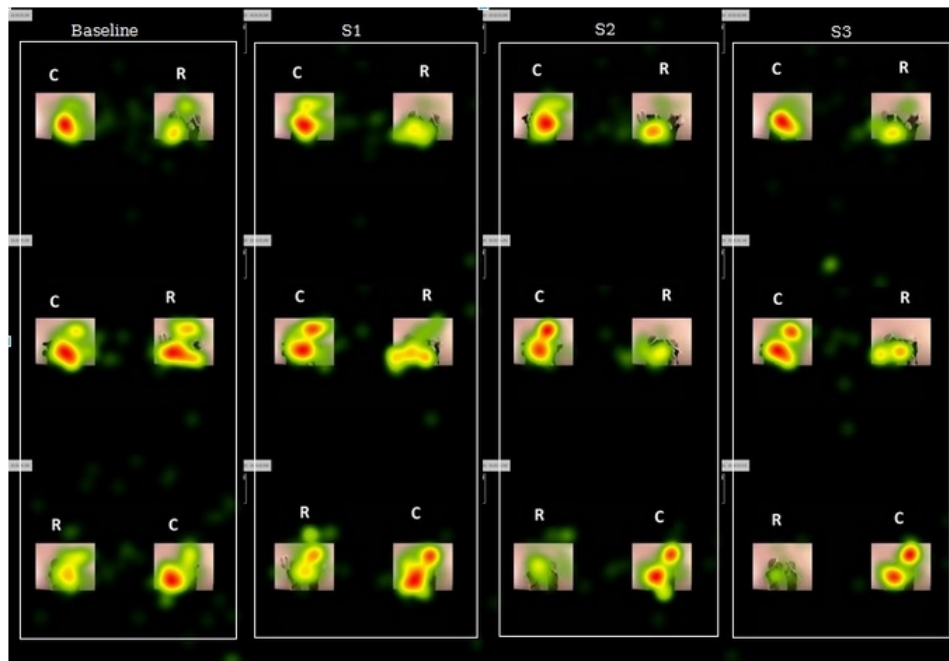
### *Coding*

The test videos were divided into four 4-second frames for analysis. As the linguistic stimuli of Hindi-Urdu was longer than French stimuli five frames were not necessary in the analysis. From the recorded videos, the duration of looking time to all test frames were computed.

#### 4.4.3 Results

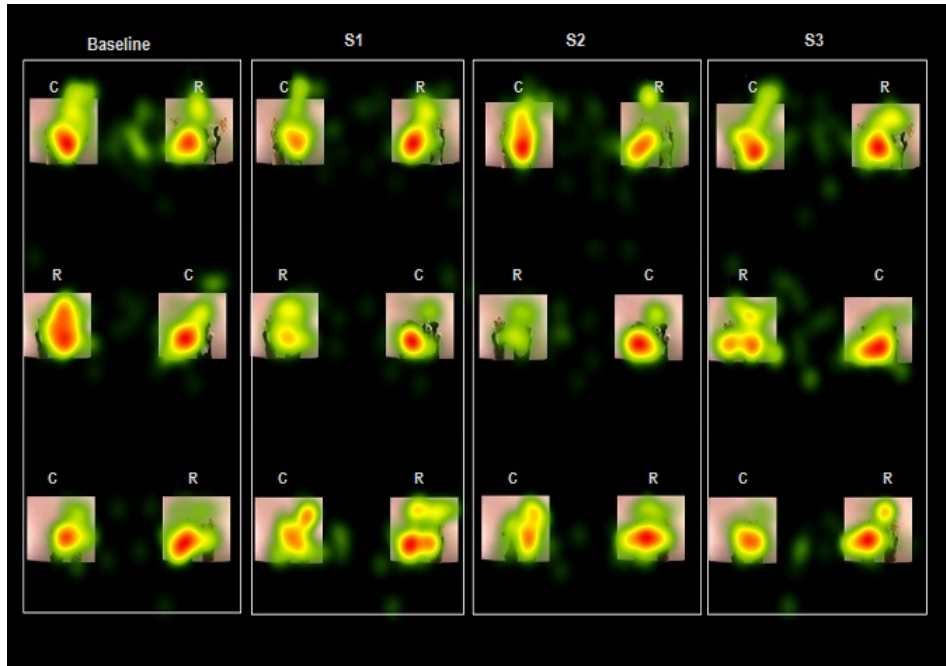
A descriptive analysis was carried out at the Psycholinguistics lab at the Facultat de Psicologia, Universitat de Barcelona. Figure 21 and Figure 22 represent the heat map of fixation duration during the grammatical and the ungrammatical conditions respectively. The four columns represent the four presentation of the linguistic input: the baseline sentence and the three presentations of the target sentence. The three rows represent the three presentations of the grammatical sentences and the three presentations of the ungrammatical sentences from the total of six presentations. The causative frames are marked with a C and the reflexive frames are marked with an R. In Figure 21, for grammatical sentences the fixation duration is longer towards the causative action which is depicted by the thicker red shade for the intensity of gaze. The performance is consistent across all the sentences and the pattern emerges in the first presentation of the test sentence itself.

Figure 21: The heat-map for grammatical sentences (C= Causative, R= Reflexive)



During the ungrammatical sentence, the children did not show any consistent performance with respect to the causative or reflexive action (see Figure 22). The intensity of gaze is fluctuating between causative and reflexive actions.

Figure 22: The heat-map for ungrammatical sentences (C= Causative, R= Reflexive)



Data were analyzed in detail at the Laboratoire de psycholinguistique expérimentale at Université de Genève. The mean looking time was computed for both causative and reflexive clips in the grammatical and the ungrammatical condition. Upon conducting a pairwise analysis on mean looking time using student t-tests, it was found that the infants looked significantly longer to the causative video than to the non-causative video only in the grammatical condition, during the first presentation of the sentence ( $t(19) = -2.549, p = .020$ ), the second presentation of the sentence ( $t(19) = -4.009, p = .001$ ), and the third presentation of the sentence ( $t(19) = -2.396, p = .027$ ). No significant difference was found in the baseline windows or for the ungrammatical sentences. (see Table 5).

Table 5: Mean looking time in ms to the causative and reflexive clips during the grammatical and ungrammatical condition

	Grammatical		Ungrammatical	
	Causative	Reflexive	Causative	Reflexive
Baseline	2250	1754	1951	1791
Sentence 1	<b>2372</b>	<b>1536</b>	1514	2072
Sentence 2	<b>2560</b>	<b>1152</b>	1785	1782
Sentence 3	<b>2021</b>	<b>1139</b>	1665	1814

From the total looking time to the causative and the reflexive clips, the proportion of looking time to the causative action was computed using Wilcoxon signed-rank test. The analysis against chance level (defined as 50%) showed above chance performance in the grammatical condition during the second presentation of the sentence window (Median proportion = 72.21;  $Z = -2.987$ ,  $p = .003$ ) as well as during the third presentation of the sentence window (Median proportion = 60.93;  $Z = -2.128$ ,  $p = .033$ ). This result is given in Figure 23 and in Table 6. None of the other windows showed above chance performance.

Figure 23: Proportion of looking time to causative action during the grammatical and ungrammatical sentences

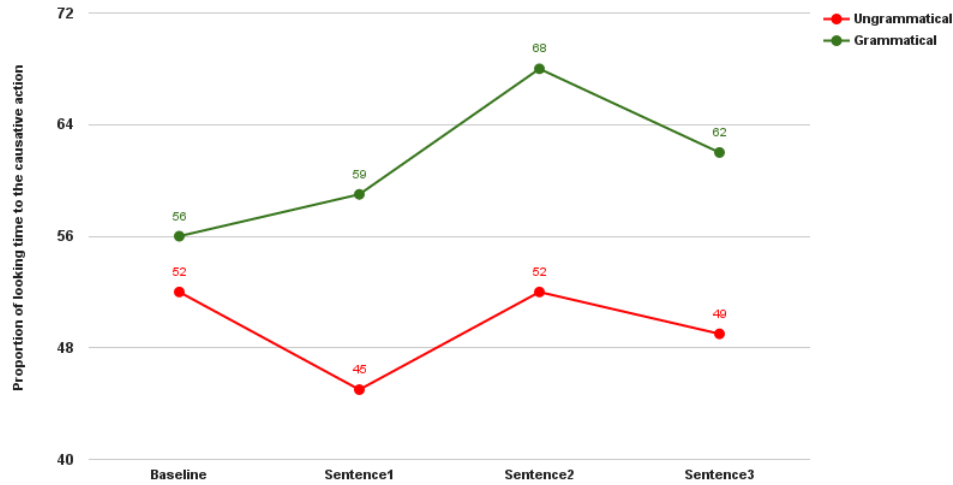


Table 6: Chance level performance in causative action

Test	chance-GmB	chance-GmS1	chance-GmS2	chance- GmS3
Z	-1,867	-1,643	-2,987	-2,128
Asymptotic significance	,062	,100	,003	,033

Test	chance-AgmB	chance-Agm1	chance-Agm2	chance-Agm3
Z	-,896	-1,157	-,560	-,149
Asymptotic significance	,370	,247	,575	,881

The comparative analysis of proportions of looking time to the causative video in the grammatical and ungrammatical sentences showed significantly higher proportions in the grammatical condition than in the ungrammatical condition during the second presentation of the sentence

(Medians : 72.21 and 52.53 respectively;  $Z = -1.979$ ,  $p = .048$ ) (see Table 7). No significant difference was found in baseline window and during the first and third presentations of the sentence.

Table 7: The comparative analysis between proportions of looking to the causative action during the grammatical and ungrammatical sentences

Test	GmB-AgmB	GmS1-AgmS1	GmS2-AgmS2	GmS3-AgmS3
Z	-,896	-1,792	-1,979	-,933
Asymptotic significance	,370	,073	,048	,351

#### 4.5 DISCUSSION AND CONCLUSION

Franck et al.'s methodology of the French experiment was replicated in Hindi-Urdu with 19-month old infants. Grammatical and ungrammatical sentences with pseudo-verbs in a video clip with animal puppets carrying out causative and synchronous reflexive actions were presented to the children. The looking time and the fixation duration was recorded using a Tobii 120 eye tracker. The children preferred to look at the causative clip for longer when they heard the grammatical transitive sentence with pseudo-verbs. No looking preference was displayed when the ungrammatical sentences were presented. The prediction that abstract word order knowledge is present in infants at very early ages was borne out.

##### *Comparison between French and Hindi-Urdu methodology*

The primary difference in the methodology between the two studies was with respect to the syntax of the sentences used. Hindi-Urdu sentences were syntactically different from French sentences in three aspects; in word order, in the use of case markers, and in the use of altered prosody as the only means to convey ungrammaticality. From the spontaneous speech analysis, the least occurring word order was chosen as all word orders in Hindi-Urdu are grammatical in different information

structure patterns. A condition where a simple clause is strictly ungrammatical was difficult to find since Hindi-Urdu allows scrambling. But in order to retain the original methodology as much as possible, an ungrammatical order in a simple transitive sentence had to be chosen. The least occurring word order VSO was chosen for this reason. So while the French experiment had SOV and SVO orders, the Hindi-Urdu experiment made use of SOV and VSO orders. This is the first syntactic difference between the sentences for Hindi-Urdu and French.

The second difference was in terms of case markers, Hindi-Urdu sentences had the agent and the patient marked with the ergative and the accusative markers. French does not make use of case markers in nominal arguments. The order of the nominals with the case markers in Hindi-Urdu was the same across grammatical and ungrammatical word orders (NP-*ne*-NP-*ko*-V and V-NP-*ne*-NP-*ko*). The use of the same order in both conditions ensured that the responses were not produced upon identifying the agent and patient role from the case markers. The question whether young infants comprehend case markers was addressed by Narasimhan & Dimroth (2008). In their study Narasimhan and Dimroth analyzed the speech production corpus of three Hindi-Urdu children for a period of one year. The analysis was to investigate the use of *-ne*, the ergative marker in Hindi-Urdu. The results showed that the young children by 15 months of age use *-ne* only in the transitive agents in perfective aspect, indicating the knowledge of the ergative marker being present in infants from early on. Thus the child grammar was found to be adult like in the use of case markers. This result adds value to the current experiment results. In the light of this result it can be claimed that even when infants in the Hindi-Urdu experiment were sensitive to the case markers they did not base their performance on case markers alone. If such was the case then the results should have shown equal preference for the causative action for both ungrammatical and grammatical condition. Both type of sentences had the agent role marked with *-ne* in the same sequence of NP-*ne*-NP-*ko*. This observation leads to the assertion that the children at 19 months did parse the sentences by virtue of the parameter setting for OV order. It is important to note here that the knowledge about the role of case markers and their sequence of occurrence did not affect the performance of infants.



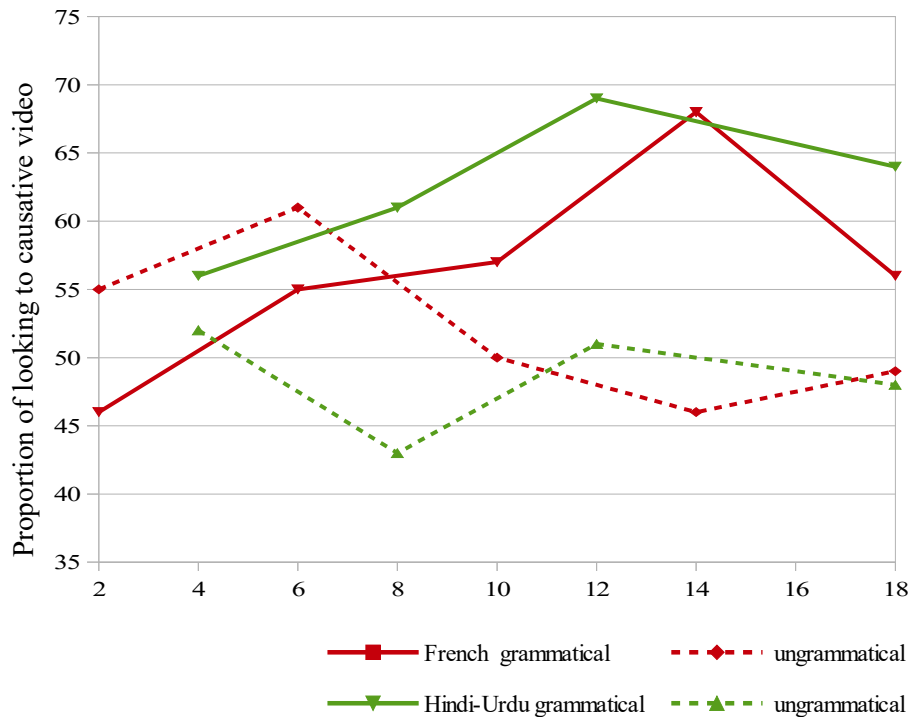
The third difference was with respect to the altered prosody of VSO order. As Kidwai (1999) argues, for a question like *What is happening?* or *What happened?* the answer is always a neutral focused SOV sentence. Non-neutral focus is used in all the other word orders where the discourse demands focus on one of the arguments or predicate. In the current experiment, the target sentences were presented after catch phrases like *Look, what is happening here?* or *Look here what happened?* were given. In this context only an answer with a neutral focus in SOV order is felicitous. An answer in wide focus intonational contour and VSO order is ungrammatical. Both SOV and VSO orders were presented in the intonation contour of a declarative sentence (Féry, 2010) which is [[L\*H] [L\*H] [H\*L]]. Since the children did not show any preference for the ungrammatical VSO it can be concluded that children are sensitive to the effect of prosody on grammaticality. In addition to the general conclusion from the results, one more conclusion can be drawn from the performance of 19-month old infants about the acquisition of prosody. It seems that the children are not only aware of the prosodic features of their native language but also the association between prosody and word order.

Regarding the morphology of the pseudo-verbs, it was pointed out that the pseudo-verb *khalayaa* gives a causative reading but *chona* gives an infinitive reading. In such a condition the morphology of pseudo-verbs could have influenced the correct performance in grammatical sentences. However, this is an incorrect generalization. In Hindi-Urdu the infinitive form of a verb appears with *-na* at the end (Kachru, 1966). So the verbs like *cheen* ('snatch'), *gin* ('count'), *maan* ('accept') which ends in *-n* in its root, appears in infinitive form as *cheen-na*, *gin-na*, *maan-na*. The transitive form of these verbs in past tense is conjugated with the perfective participle *-a*, resulting in *cheena*, *gina*, *maana*. According to this descriptive rule, the infinitive form of the pseudo-verb *chon* would be *chon-na* and the transitive form *chona*. So the morphology of the pseudo-verbs did not interfere in giving a causative reading.

In the French experiment the data was analyzed by dividing each test video into 5 frames, but for Hindi-Urdu only four frames could be accommodated based on the start and end of the linguistic stimuli. The Hindi-Urdu sentences were longer than the French sentences. This

presents a difficulty in comparing the French and Hindi-Urdu results on a frame by frame basis. However the progression of performance from baseline sentence to the third presentation of the test sentence can be compared. In Figure 24, the proportion of looking time for both ungrammatical and grammatical conditions for Hindi-Urdu and French is illustrated. Despite the difference in timing infants in both studies seem to be following a similar pattern in performance. When the grammatical sentence was heard for the first time children in both studies showed an increment in looking time towards the causative clip, with further increment during the second presentation of the grammatical sentence. By the third presentation there was a decrease in looking time, which could be due to saturation effect.

Figure 24: Comparison between French and Hindi-Urdu results



When Franck et al.'s methodology was replicated in Japanese, it was found that the children displayed sensitivity to head-final order only by 30 months of age. This delay in acquisition was proposed to be due to the presence of null arguments in Japanese. But Hindi-Urdu also has null arguments and Hindi-Urdu children did show an early knowledge of the head-final order. It is unclear why Japanese children did not show any knowledge of the head-final order in early ages. However, evidence from Hindi-Urdu is conclusive on the fact that despite the presence of null arguments and free word order children have set the parameters for the OV order at 19 months of age.

A criticism that was raised on the conclusions of the Hindi-Urdu acquisition result was that the results were only indicative of how children performed when the OV order was presented. The criticism mainly pointed out that the performance cannot measure up to claiming a parameter is set, but it is only suggestive of the sensitivity of the children to word order. However, if children had not set the parameter, on what knowledge they based their performance becomes a puzzle. I conclude that the fact that children parsed grammatical SOV in the presence of a pseudo-verb shows that an intrinsic system which could identify grammaticality from the word order is already functioning in 19-month old infants.

#### 4.5.1 *Usage based or Universal grammar?*

In the light of the French and Hindi-Urdu results, it can be concluded that children do not acquire word order as in the lexical learning theory. The lexical learning theory and the Verb-Island hypothesis (Tomasello, 1992, 2000) fail to account for the performance of infants from both Hindi-Urdu and French experiments. If children were acquiring word order by memorizing the position of lexical items of each verb as templates, then parsing of word order information from a sentence with a pseudo-verb would be impossible. But the empirical evidence from the current experiment suggests otherwise. The Usage-based accounts fail especially in the light of the fact that there was no familiarization with the pseudo-verbs anywhere in the procedure. So the correct parsing of the OV order as grammatical and thereby attaching causality to it from

a transitive OV sentence indicates that there is abstract knowledge of OV order present in the young infants by 19 months of age. Use of pseudo-verbs also refutes the argument that children learn word order as a verb-specific knowledge from the more frequently occurring verbs (Matthews et al., 2007).

Dittmar et al. (2008) raised a criticism against preferential looking studies which employed training phases in experimental design (Gertner et al., 2006). It was pointed out that, from the videos and sentences presented during the training phase the children had the possibility to extract a generalization that the NP that appears at the sentence-initial position always corresponded to the agent role. In Gertner et al.'s study sentences like (106) were presented with two clips where the duck was the agent in both target and distractor clip.

(106) The duck is hugging the bunny.

The same nominals were used in test trials. So Dittmar et al. point out that the construction, *the duck - verb- the bunny*, could have resulted in a priming effect. In the Hindi-Urdu experiment at no point in the familiarization phase nominals used in the test sentences were provided in a transitive form. It is impossible to draw a conclusion that children had a chance to learn the agent and patient roles from familiarization videos. Moreover, even if the children did pick up the clues about the agent and the patient roles, maybe from case markers, both the grammatical and the ungrammatical sentences were presented with the same NP-NP sequence with the same case markers. This would have resulted in children parsing the ungrammatical sentence and thereby looking at the causative action. The results from Hindi-Urdu experiment show otherwise.

Hindi-Urdu and French results invalidate proposals (Akhtar, 1999; Tomasello, 2000; Abbot Smith et al., 2001) about children possessing only lexical specific knowledge about word order properties from the Usage-based theory.

*Why chance performance for VSO order?*

The fact that the children did not assign a causative or reflexive interpretation to the ungrammatical VSO sentence has to be discussed. The children parsed the SOV sentence with a pseudo-verb and assigned causality to it. This indicated the knowledge of the abstract head-final order. So it is worthwhile to look into what happened when VSO sentences were given.

About the comprehension of prosody, literature is abundant with studies which prove that young infants are sensitive to prosody of their native language (by 6-12 weeks, according to [Christophe et al. \(2003\)](#)), and the prosody of the language specific clause-internal phrase boundary (by 9 months according to [Juszyk et al. \(1993\)](#)). For an infant acquiring a language which codes information structure with a combination of prosody and word order, this sensitivity is more crucial. The fact that the wrong pairing of prosody and word order results in ungrammaticality is a rule a young language learner must master. There is a three fold task here for acquisition. One: they must isolate from a data of mixed word orders which is the unmarked basic word order. Two: they must accept the marked word orders to be grammatical by identifying which prosody is associated with different structural configuration. Three: discourse contexts of marked and unmarked orders should be differentiated. In terms of parameter acquisition of a free word order language like Hindi-Urdu, it becomes logical to argue that the children must first have some knowledge about the head-final order prior to acquiring variations within word orders. As it was pointed out in the hierarchy of parameter acquisition by [Roberts \(2012\)](#) for the head-directionality parameter, first the system analyzes if there is a head-final feature present in all heads. After this macroparametric level analysis and setting of the OV parameter, the child must establish the more marked word orders. If Robert's argument is accepted then the parameter for the head-final order is set very early in Hindi-Urdu acquirers. So after setting the macro parameter for the head-final order, the system, upon encountering a set of marked word orders, would be guided to set the further word order parameters. More clarity on this can be obtained from the under performance of the subjects on the ungrammatical sentences in the

Hindi-Urdu experiment. There could be two possibilities through which the ungrammaticality must have been assigned. Either children had no knowledge of the marked word orders and thereby did not parse VO order, or children were sensitive to the marked word orders and their felicitous discourse contexts, thereby assigning ungrammaticality to the infelicitous VO order. The exact reason is difficult to pinpoint at this stage. Further investigations are due in order to reach a conclusion.

Part IV

THE ACQUISITION OF MALAYALAM WORD  
ORDER

## THE ACQUISITION OF SCRAMBLING

---

### 5.1 INTRODUCTION

The purpose of this chapter is to report the study on the acquisition of scrambling in Malayalam. The previous section provided evidence from an Indo-Aryan language about the acquisition of the basic OV order. I argued for the early setting of the parameters responsible for the OV order, and for adopting the UG based theory as the best explanation for the logical problem of language acquisition. In this study, the same assumptions are retained, and the attempt is to investigate the early acquisition of scrambling. The chapter proceeds as follows. First, a review of the scrambling acquisition literature is given, followed by sections on the experimental hypothesis, methodology, experiment and the results, and finally the discussion of the data.

### 5.2 THE LITERATURE ON THE ACQUISITION OF SCRAMBLING

Scrambling is found to be operative in child grammar as early as the two-word production stage in languages like Japanese (Sugisaki, 2008), Dutch (Schaeffer, 2000a), Russian (Avrutin & Brun, 2001), Turkish (Batman-Ratyosyan & Stromswold, 1999), and Ukrainian (Mykhaylyk, 2012). The studies reported on the acquisition of scrambling in Japanese occupy a core place in the literature of scrambling research. Japanese scrambling experimental results are reported from as early as the sixties. The early accounts (Hakuta, 1977; Sano, 1977) provided results that contradicted the idea that scrambling is acquired early. The reason for this finding was attributed by others to the fact that scrambled sentences were provided without discourse cues in the studies. Contrary to the early accounts, later studies (Otsu, 1994; Murasugi & Kawamura, 2005) investigated scrambling with discourse cues and found that the acquisi-



tion of scrambling is early in Japanese. Some of the relevant studies are given in detail in the following section.

*Scrambling with no discourse context*

Hakuta (1977) researched upon the role played by word order and case markers to parse sentences in Japanese child grammar. Hakuta assumed the perceptual comprehension strategy that Bever (1970) proposed for English, that is, for any NP-NP-V sequence children employ an agent-patient-action configuration while parsing sentences. This strategy explains that children would assign thematic roles based on the order of the DPs in the surface order. Hakuta claimed that if children used word order information for processing sentences employing Bever's strategy, then there would be a difference in performance with active SOV and passive OSV sentences. But if it was the case that particles were used to process a sentence then children should not face any difficulty in any of the active or passive word orders, since they all contain the case markers indicating agent and patient. Four types of word orders were employed in the test sentences: SOV active (107 a), OSV active (107 b), SOV passive (107 c), OSV passive (107 d).

- (107) (a) *Giraffe-ga Tiger-o licked.*  
giraffe-NOM tiger-objective-particle lick-PAST  
'The giraffe licked the tiger.'
- (b) *Tiger-o Giraffe-ga licked.*  
Tiger-objective-particle giraffe-NOM lick-PAST  
'The giraffe licked the tiger.'
- (c) *Tiger-ga Giraffe-ni licked.*  
tiger-NOM giraffe-object lick-PAST-PASSIVE  
'The giraffe was licked by the tiger.'
- (d) *Giraffe-ni Tiger-ga licked.*  
giraffe-object tiger-NOM lick-PAST-PASSIVE  
'The giraffe was licked by the tiger.'

The verbs used were *ketta* ('kicked'), *nameta* ('licked'), *butta* ('hit'), *kisushita* ('kissed'), and *kusugutta* ('tickled'). An act-out task was carried

out to test comprehension, on subjects who were divided into four groups based on the MLU. The mean age of each group was: 3;6 years (Group I), 3;9 years (Group II), 5;1 years (Group III) and 5;4 years (Group IV). Following the comprehension test, a production task was also given. Children were shown slides depicting various actions using a slide viewer and were asked to describe what they saw in the slides to a puppet. Responses from the act out task was scored as either correct or wrong, and among the wrong responses, role reversals were noted. The presence of case markers and the order of subject and object were analyzed from the production data.

Group I, with the lowest mean age, performed above chance for the SOV actives and below chance for the OSV passives. Since both these orders employed agent-patient-action order, it was concluded that younger children did not rely on word order alone to parse sentences. Children were found to be performing below chance for OSV active and SOV passive sentences. These sentences involved the patient-agent-action order and children consistently interpreted them in reversed roles. Above-chance performance for active sentences in both orders was shown only by children in Group III and Group IV; that is by children from the age of 5-years and above. In general, children reversed roles for passive SOV sentences and performed accurately for active SOV sentences. It was concluded that children were employing Bever's strategy with an addition, that is, the first agent marked with *-ga* was taken to be the agent. Every first noun was not taken as the agent; if that was the case children could have performed better for OSV passive where the first noun was the agent. Hakuta claimed that since the agent in the OSV passive was marked with *-ni*, in the absence of *-ga* marker children dismissed the first noun as the agent. So a certain correlation between case markers and word order was proposed about how child grammar interprets sentences. About the acquisition of scrambling, since only the children in group III with the mean age of 5;1 did perform above chance for active OSV sentences, that lead to the conclusion that scrambling acquisition is delayed in Japanese. In the production data out of the 1200 sentences analyzed only three instances of the non-canonical order were noted.

The results from Hakuta's study elicit a conclusion that both comprehension and production of scrambled sentences are delayed in Japanese. But this result was not replicated in later studies. For instance, more occurrences of scrambled orders in the spontaneous production of young Japanese children were found by Iwatate (1981). Spontaneous production data of five 2-3-year-old children were analyzed by Iwatate. A 48 hour recorded speech corpus was examined from each subject, and sentences in SOV, OSV, SVO and OVS word orders were attested. Another speech production data analysis by Sugisaki & Otsu (2011) reported that Japanese children by the age of 2.5 years produced VO utterances like (108) which also proves the presence of scrambling in young ages.

- (108) *Morattekita, kore.*  
           got               this  
       '(I) got this.'

*Scrambling with discourse context*

Otsu (1994) argued that scrambled sentences out of context might sound unnatural and thus they might hinder the sentences from being interpreted correctly. Otsu inserted contextual cues to the scrambled sentences to test the of acquisition of scrambling. The hypothesis was that children by the age of three would have acquired scrambling. As shown in (109) sentences in this study had a discourse context.

- (109) *Kooen ni ahiru-san ga imasita, Sono ahiru-san o*  
       park in duck    NOM is-POL-PAST, the   duck   ACC  
       *kame-san ga osimasita.*  
       turtle    NOM push-POL PAST  
       'There was a duck in a park. A turtle pushed the duck.'

There were two groups in the study, one control group (six 3-year-olds, and six 4-year-olds) who received no context and one experimental group (six 3-year-olds, and six 4-year-olds) who received contextual cues. Both groups received a practice session with two transitive and two intransitive sentences. The control group received sentences as (110) and experimental group received sentences as (109). Children were asked to act out the sentences.

- (110) *Ahirusan-o kamesan-ga osimasita.*  
 duck-ACC turtle-NOM push-PAST  
 'A/The turtle pushed a/the duck.'

Results revealed that the experimental group interpreted scrambled sentences correctly for 90%, whereas the control group only scored 46%. The performance of children based on their age is given in Table 8. This result favored the early acquisition of scrambling proposal.

Table 8: Correct responses by experimental and control group (Otsu, 1994)

	Experimental group	Control group
3 years	42%	12.5%
4 years	48%	33%

Sugisaki & Isobe (2001) sought evidence from the acquisition data for the movement proposals of scrambling. While it is conclusive that M-scrambling involves A' movement, S-scrambling has been explained with two distinct proposals. One that says DO-IO results due to the movement from IO-DO (Tada, 1993), the other that proposes that both the orders are base generated (Miyagawa, 1997). Sugisaki and Isobe predicted that if the Japanese children had more difficulty comprehending S-DO-IO-V than S-IO-DO-V, then the derivation of DO-IO is via A-movement. If the difficulty is the same for both the orders, then that would support the proposal that both orders are base generated. If children performed better with sentences involving M-scrambling and poor with S-scrambling then two conclusions were argued to follow: S-scrambling is the result of movement, and it is A-movement that results in S-scrambling. The primary hypothesis was that Japanese children would have more difficulty in scrambling that involves A-movement. The hypothesis was developed on Tada's observations (Tada, 1993) that S-scrambling involves A-movement and M-scrambling involves A' movement. A-movement has been extensively argued to be problematic for children (Borer & Wexler, 1992). Twenty children from the age range of 3;11 to 5;0 years were tested. A story was narrated, after which a character described what happened in the story, and children were asked

for a truth value judgment. A total of nine sentences were tested, two canonical order sentences (SUB-IO-DO-V (111, 112)), one S-scrambled sentence (SUB-DO-IO-V (113)), one M-scrambled sentence (IO-S-DO-V (114)) and five filler sentences for the task.

- (111) *Satoshi-ga akachan-ni Pokemon-o misetayo.*  
 Satoshi-NOM baby-DAT Pokemon-ACC showed  
 'Satoshi showed his Pokemon to the baby.'

- (112) *Ookido Hakase-ga Kasumi-ni atarashii Pokemon-o*  
 Dr. Ookido-NOM Kasumi-DAT new Pokemon-ACC  
*misetayo.*  
 showed  
 'Dr. Ookido showed a new Pokemon to Kasumi.'

- (113) *Satoshi-ga Pikachu-o okaasan-ni misetayo.*  
 Sathoshi-NOM Pikachu-ACC mother-DAT showed  
 'Sathoshi showed Pikachu to his mother.'

- (114) *Kasumi-ni Satoshi-ga Pichu-o misetayo.*  
 Kasumi-DAT Satoshi-NOM Pichu-ACC showed  
 'Satoshi showed Pichu to Kasumi.'

The results showed that children performed above chance for sentences involving M-scrambling (90%), and canonical order (85%, 90%). This finding lead to the conclusion that Japanese speaking children acquired M-scrambling, which is a result of A' movement, as early as 3 years. Sugisaki and Isobe proposed that the knowledge of M-scrambling could be present in child grammar even before the age of three years. Regarding the S-scrambling, it was found that children performed poorly in sentences involving S-scrambling. Only 60% of correct responses were obtained for DO-IO order. The fact that children had more difficulty in DO-IO order when compared to IO-DO order lead the authors to claim that the S-scrambling in Japanese involves A-movement. In the light of these results, it was suggested that since scrambling involves different movement operations, the acquisition of different scrambled orders would show a difference in the age of acquisition.

Murasugi & Kawamura (2005) investigated the age of acquisition of scrambling and whether or not children possess the knowledge of the

reconstruction property of scrambling. First, the age of acquisition was tested using active sentences (115), passives (116), and scrambled sentences (117).

- (115) *Ahiru-ga usi-o oikake-ta.*  
 duck-NOM cow-ACC chase-PAST  
 'The duck chased the cow.'

- (116) *Usi-ga ahiru-ni oikake-rare-ta.*  
 cow-NOM duck-by chase-passive-past  
 'The cow was chased by the duck.'

- (117) *Usi-o ahiru-ga oikake-ta.*  
 cow-ACC duck-NOM chase-PAST  
 'The cow, the duck chased.'

Twenty-one test sentences, seven from each type, were tested, and the subjects were asked to act out the sentences. If children interpreted the right agent and patient, that response was coded as correct. Twenty-two children within the age range of 2-6 years were tested. Individual performance by all the subjects is given in Table 9.

Table 9: Percentage of correct response (Murasugi and Kawamura, 2005)

Subject	Age	ExpI Active %	Exp I Scram- bling %	Exp I Passive %	Exp II Active %	Exp II Scram- bling %	Exp III Passive %
A	2	83	83	50	0	NT	NT
B	2	83	66	17	0	NT	NT
C	3	100	100	100	100	100	50
D	3	100	100	28	100	100	33
E	3	100	100	42	100	100	16
F	3	28	42	0	50	38	16
G	3	71	71	28	66	50	50
H	3	100	85	57	83	87	50
I	4	100	100	0	100	100	33
J	4	100	100	71	100	100	33
K	4	100	100	42	66	75	16
L	4	100	100	85	83	87	33
M	4	100	100	100	100	100	33
N	4	100	100	100	100	100	50
O	5	100	100	100	100	100	100
P	5	100	100	100	100	100	100
Q	5	100	100	100	100	100	100
R	5	100	100	100	100	100	100
S	5	100	100	100	100	100	33
T	5	100	100	100	100	100	50
U	6	100	100	100	100	100	100
V	6	100	100	100	100	100	50

The two 2-year-old children in the study interpreted scrambled sentences for 83% and 66% of the time, and three out of six 3-year-old children for 100% of the time. The high percentage of correct responses

with scrambled sentences from the age of 2 was in support of the fact that scrambling is acquired early.

A second experiment was conducted to check if children are aware of the reconstruction property of scrambling. This was checked using the anaphor *zibun*, which is subject oriented. Anaphors need a c-commanding subject antecedent and when the sentence is scrambled this relation does not hold any longer. However, the sentence can be interpreted the same as a non-scrambled sentence after reconstruction. In this experiment children were given twenty sentences with active (115), scrambled (119) and passive (120) constructions with the anaphor *zibun*.

- (118) *Ahiru-ga<sub>i</sub> usi-o [zibun-no niwa-de]<sub>i</sub> oikaketa.*  
 duck-NOM cow-ACC self-GEN garden-at chased  
 'The duck<sub>i</sub> chased the cow at the garden of himself<sub>i</sub>.'
- (119) *Usi-o [zibun-no niwa-de]<sub>i</sub> ahiru-ga<sub>i</sub> oikaketa.*  
 cow-ACC self-GEN garden-at duck-NOM chased  
 'The cow, at the garden of himself, the duck chased.'
- (120) *[Kuma-ga]<sub>i</sub> usagi-ni [zibun-no niwa-de]<sub>i</sub> oikakerareta.*  
 bear-Nom rabbit-by self-Gen garden-at was-chased  
 'The bear<sub>i</sub> was chased by the rabbit at the garden of himself<sub>i</sub>.'

The subjects, task, and coding were the same as in Experiment I. The results are given in Table 9. It was found that children at the age of 2 could not parse the SOV active sentence with the anaphor, which led to the conclusion that they did not have the lexical knowledge of the anaphor *zibun*. The older children interpreted the sentences with the anaphor in the adult like manner which proved that children at 3 possess the reconstruction property of scrambling. Apart from the results with the anaphor, the fact that the 2-year-old children performed at high percentages for scrambled sentences shows that scrambling is available at the age of two. The poor performance was attributed to the presence of the anaphor and not due to inability to parse scrambled sentences. It was also noted that scrambling is acquired early than passives, as children who had good performance with scrambled sentences had difficulty with passives.



*Scrambling studies in other languages*

Schaeffer (2000b) reported that child Dutch exhibits optional scrambling as opposed to the obligatory scrambling in adult Dutch. In adult grammar, a direct object can appear before or after negation and adverbs in the space between C and the finite verb. The motivation for object scrambling across negation and high adverbs (the adverbs that occur at TP level) is the feature [referential]. Referentiality is a semantic notion marked with the syntactic feature [referential]. A referential noun is a unique member of a particular class which has a fixed referent in the world, and a non-referential noun is any member of that class. Furthermore, in a discourse, the [referential] feature can be applied to an object that has already appeared prior to the target sentence; this is termed discourse related referentiality. One of the claims of the study, which is relevant to the acquisition of scrambling, was that children would fail to scramble object DPs as they lack the ability to mark referentiality. Schaeffer claimed that optional scrambling is found in child grammar because there is a delay in distinguishing between discourse related and non-discourse related referentiality. Her experiment was a combination of Truth Value Judgment Task and Elicited Production Task. One of the experimenter described a scene and an assistant who was playing the role of a silly tiger said something at the end which was not true, and the child had to say if it was true or not and if not true the child was prompted to correct the silly tiger. The two answers by children in the test scenario involving adverbs are given in (121), the first with object scrambling and adult-like.

(121) Scene: picture of a tree

Cookiemonster: *-Kijk, een boom. Die vind ik zo mooi, die ga ik MOOI inkleuren.*

Tom the Tiger: *-De boom gaat Koekiemonster LELIJK inkleuren!*

Child: *-Nee!*

Experimenter: *-Nee he? Wat gebeurt er echt?*

Cookiemonster: *-Look, a tree. I find it so beautiful, I'm going to color it BEAUTIFULLY.*

Tom the Tiger: -Cookiemonster is going to color the tree  
IN-AN-UGLY-WAY!

Child: -No!

Experimenter: -No? What's really happening?

(122) EXPECTED RESPONSE :

*Koekiemonster gaat de book MOOI inkleuren!*  
Cookiemonster goes the tree beautifully in-color  
'Cookiemonster is going to color the tree BEAUTIFULLY'

As the prediction was that children would fail to obligatorily scramble the object DP, a response with no scrambling (123) was also possible.

(123) *Koekiemonster gaat MOOI de book inkleuren!*  
Cookiemonster goes beautifully the tree in-color  
'Cookiemonster is going to color the tree BEAUTIFULLY'

A total of 45 sentences were tested, which involved five types of object DPs (definite DP, personal pronoun, referential indefinite, nonreferential indefinite, proper name) over three environments in which objects obligatorily scramble (negation, high adverb, low adverb). Each object item was tested in each environment for three tokens. In line with the prediction, in (121) children would place the definite DP after the adverb, contrary to adult grammar which requires obligatory scrambling of object to a position preceding adverb (/negation)

Since 2-year-old children do not use adverbs frequently, the data was analyzed separately on their performance in sentences involving negation. Only this result is discussed here. It was found that 2-year-old children scrambled definite object DPs and proper names to pre-negation position only for 30% and 31% of the time respectively. But the performance of the 3-year-old children differed significantly from the 2-year-olds; they scrambled objects to the pre-negation position for 72% (definite DP) and 73% (proper name). There was no significant difference found between the performance of the 3-year-old children and the adults, which indicated that by the age of three children develop the pragmatics and syntax of scrambling. The prediction that young children would fail to scramble objects obligatorily was borne out.

In addition to Japanese and Dutch, studies from languages like Russian, Serbo-Croatian, Ukrainian and Turkish also report early acquisition of scrambling (Batman-Ratyosyan & Stromswold, 1999; Avrutin & Brun, 2001; Ilic & Deen, 2004; Dyakonova, 2005; Mykhaylyk, 2012). Evidence from acquisition literature is conclusive on the fact that scrambling is acquired early. However in all of the above studies the properties of scrambling investigated are divergent. In Dutch, acquisition of object scrambling due to a referentiality feature was in focus, in Ukrainian, Russian and Serbo-Croatian acquisition of scrambling associated with specificity and definiteness was studied, in Japanese movement types of scrambling, and scrambling with and without discourse cues were studied. As the motivation for scrambling and the constraints on scrambling differ cross-linguistically (Bošković & Takahashi, 1998; Grewendorf & Sabel, 1999) it becomes justifiable to investigate scrambling acquisition cross-linguistically. The data from Malayalam is expected to add information about how child grammar deals with clause internal A' movement of arguments under information structure constraints. The experiments reported in the subsequent sections detail the investigation on the acquisition of clause internal scrambling in Malayalam.

### 5.3 EXPERIMENT I

#### 5.3.1 *Rationale*

Though scrambling is necessitated by discourse, adult speakers of Malayalam do comprehend scrambled sentences without any discourse context. To claim that once the parameters of word order are set, the knowledge of scrambling in child grammar is the same as adult grammar, it is essential to match the linguistic performance of adults and children in the same controlled environments. The current experiment was designed with a methodology that controlled the discourse context variable. The early acquisition studies (Hayashibe, 1975; Hakuta, 1977) that investigated the acquisition of scrambling without discourse contexts, in Japanese, served as the background for this experiment. Taking the Japanese studies and the fact that the adults perform well in Malayalam scrambling without discourse context, the aim of the current experiment was formulated, which was to investigate if child grammar parses scrambled sentences without a discourse context. My prediction for this experiment was that young children by the age of 3 would perform adult like in interpreting scrambled sentences without discourse context. This was based on the hypothesis of VEPS, which was that the parameters that determine the clause internal middle scrambling involving A' movement in Malayalam are acquired early.

#### 5.3.2 *Experimental Design*

The experiment designed was a picture matching task.

##### 5.3.2.1 *Subjects*

A total of 60 children and 20 adults participated in the experiment. The adults served as the control group. Both the adults and the children were native speakers of Malayalam. All the children were students of the Ananthapuri School at Trivandrum, Kerala. The experiment was conducted at the school. Based on the mean age, the child subjects were divided into three groups (see Table 10). From the 60 children, 28 were

males, and 32 were females. From the 20 adults, 8 were females, and 12 were males. Adult subjects were within the age range of 21 - 32 years (mean age 24;4 years).

Table 10: Subjects

	Age range	Mean age	Number
Group I	3;2 - 3;11	3;5	20
Group II	4;3 - 4;9	4;3	20
Group III	5;0 - 5;5	5;3	20

### 5.3.2.2 *Materials*

A set of pictures with gender matched characters in each picture were used (see Figure 25). Every picture contained four frames: one frame in which the same agent and patient as the target word order carrying out the action, a second frame with the agent-patient roles reversed, a third frame where a different agent was performing the action and a fourth frame where no characters were performing any action. All the four frames were randomly arranged in each picture. A laptop was used to present the sentences in a slide show.

Figure 25: A sample of the picture used in Experiment I



### Test Sentences

Four types of word orders were chosen: the basic word order, SOV (124), three middle scrambled sentences of the order OSV (125), SVO (126) and OVS (127). Six tokens of each word order were presented, each child was tested on a total of 24 sentences. The subject arguments were marked with the nominative null marker, and the objects were overtly marked with either an accusative marker or a dative marker.

- (124) *aniyathi*                      *chechiy-e*                      *thallunnu*.  
 younger sister-NOM elder sister-ACC push-PRES  
 'The younger sister is pushing the elder sister.'

- (125) *muthashiy-e*                      *aniyathi*                      *thodunnu*.  
 grandmother-ACC younger sister-NOM touch-PRES  
 'The younger sister is touching the grandmother.'

- (126) *achan puthappikkunnu aniyam-e.*  
 father-NOM cover-PRES younger brother-ACC  
 'The father is covering the younger brother.'
- (127) *aniyane orukkunnu achan.*  
 younger brother-ACC dress-PRES father-NOM  
 'The father is dressing the younger brother.'

A total of eleven transitive verbs (one verb was repeated twice in both sets of pictures) were used in the test sentences. The root form of all the verbs used are given in Table 11. The verbs were tense marked for present progressive with the particle *-unnu* in the test sentences. Not all verbs were used in all the four word orders. The verb-final orders (SOV and OSV) were presented with one set of verbs, and the verb medial orders (SVO and OVS) were presented with another set of verbs.

Table 11: Verbs used in Experiment I

verbs used in SOV and OSV order	verbs used in SVO and OVS order
<i>odikkuka</i> 'chase'	<i>puthappikkuka</i> 'cover'
<i>ummaavekkuka</i> 'kiss'	<i>sahaayikkuka</i> 'help'
<i>orukkuka</i> 'dress'	<i>adikkuka</i> 'hit'
<i>vaarikodukkuka</i> 'feed'	<i>kettippidikkuka</i> 'hug'
<i>thalluka</i> 'push'	<i>thudakkuka</i> 'wipe'
<i>thotuka</i> 'touch'	<i>thotuka</i> 'touch'

### 5.3.2.3 Procedure

The subjects were tested individually in a separate classroom in the school. To develop a rapport with the children first the experimenter and assistant were introduced by the teachers and there was a short interaction session in the classroom. Children were told by their teachers that the new ladies would show them colorful pictures on a laptop. After the interaction session in the classroom, once the set up was arranged, children were sent one by one to the experiment room. Initially,

the materials were shown to the children for a familiarization session. The experimenter showed independent pictures of characters and introduced them. Each child was made to point to the characters when the experimenter asked *which is the grandmother, which is the younger sister*, to ensure that the drawings in the picture were not confusing. After this the children were given the following instruction:

(128) Experimenter: -*“Now a screen with four pictures will appear, and I will say a sentence out loud. After hearing the sentence you have to either point to or touch the matching picture on the screen. If you don’t hear the sentence properly, ask again and I will tell you. I will wait until you point to a picture to say the next sentence.”*

Schmitt & Miller (2010) support using the Picture-Matching task with young children on the observation that complex sentences which can be depicted by pictures can be tested easily. When children are given different choices for the possible interpretations of a sentence, they are forced to choose the one picture that most approximates their interpretation of the sentence. Schmitt and Miller stress the importance on providing enough time to decide on which picture to choose so that it can be made sure that the children considered all the possible options. This was followed in the experimental procedure.

The research assistant who recorded the responses was standing behind the subjects but facing the screen. Before presenting the test sentences, for each picture the experimenter caught the attention of the child by saying *look here, what are they doing?* No reinforcement or feedback was provided until the end of the experiment. Each child took about fifteen minutes for the entire experiment. All the adults were tested in a home environment using the same materials and procedure.

### *Coding*

If the children pointed to the picture with the same agent and patient as the target word order, then the answer was scored as C (for correct). If they pointed to the agent-patient roles reversed picture then it was scored as R (for reverse). If the children pointed to the picture with a different agent than that from the test sentence then it was scored as O



(for other agent) and if they pointed to the picture where no character was performing any action it was scored as N (for none).

### 5.3.3 *Results*

All the adults performed 100% for all word orders. For the data obtained from children, a descriptive analysis was carried out first, and then modeling of the proportion of correct answers with respect to age and sentence type were calculated.

#### 5.3.3.1 *Descriptive analysis*

The results were analyzed based on correct and incorrect responses. Every response (reverse, other and none) except correct was included in the incorrect count. Table 12 gives the percentages and the counts of correct and incorrect answers for all word orders for each age group.

Table 12: Raw score and percentage of correct and incorrect responses

Group (mean age)	Order	Correct		Incorrect	
		Count	%	Count	%
Group I (3;5 years)	SOV	83	69%	37	31%
	OSV	86	71.7%	34	28.3%
	SVO	86	71.7%	34	28.3%
	OVS	70	58.3%	50	41.7%
Group II (4;3 years)	SOV	87	72.5%	33	27.5%
	OSV	64	53.3%	56	46.7%
	SVO	76	63.3%	44	36.7%
	OVS	77	64.2%	43	35.8%
Group III (5;3 years)	SOV	105	87.5%	15	12.5%
	OSV	95	79.2%	25	20.8%
	SVO	109	90.8%	11	9.2%
	OVS	106	88.3%	14	11.7%

Figure 26: Results for Group I (mean age - 3;5 years)

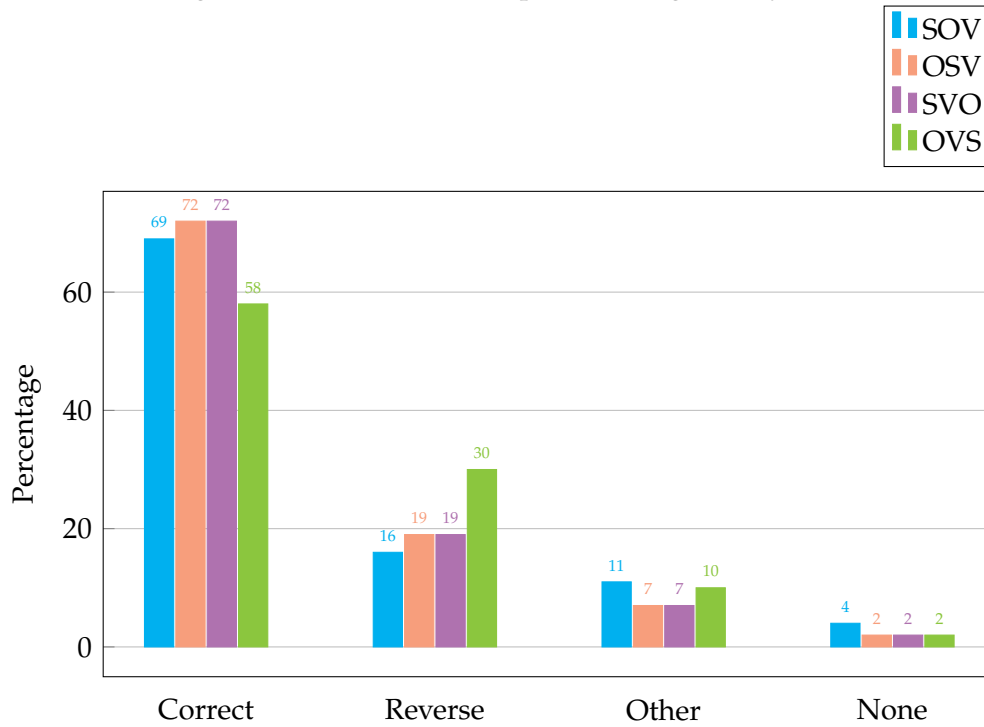


Figure 26 presents the percentage of responses by the 3-year-old children across all word orders. The percentage of correct answers for all word orders are above 50%. The maximum number of reverse answers by the 3-year-old children were found in the OVS order (30%).

Figure 27: Results for Group II (mean age - 4;3years)

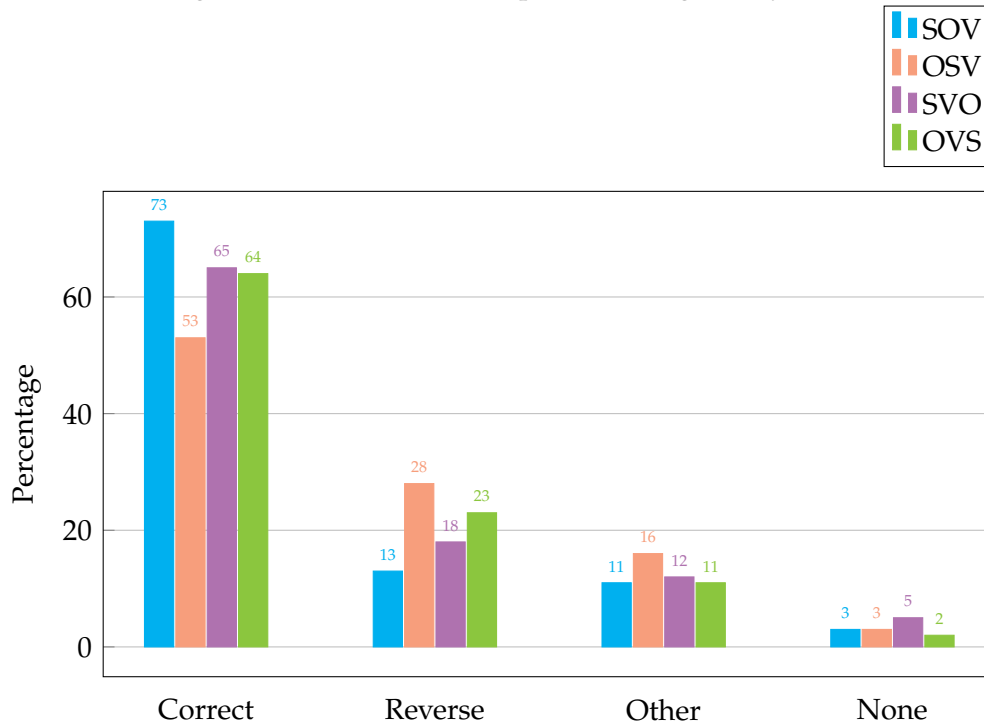
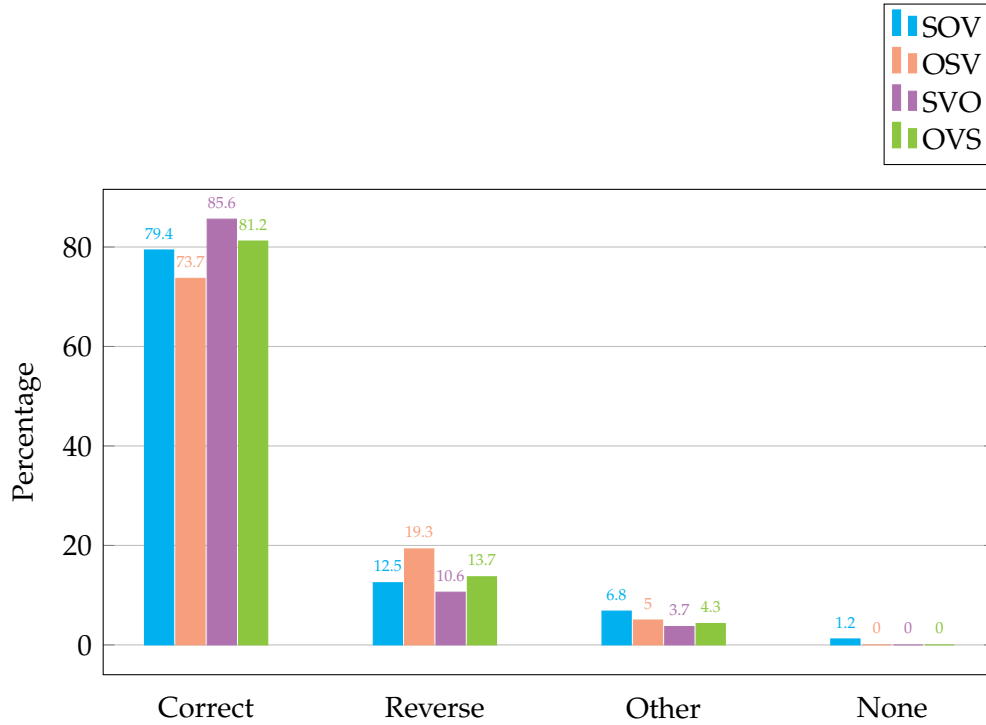


Figure 27 shows the performance of 4-year-old children on all word orders. This age group showed an overall reduced performance when compared to the other age groups. However, all the word orders were correctly interpreted for more than 50% of the time. The order OSV was found to have the least number of correct answers (53%). The errors that the 4-year-old children made with the OSV order were more towards reversal (28%), indicating that the 4-year-old children had difficulty interpreting agent and patient when the order was scrambled.

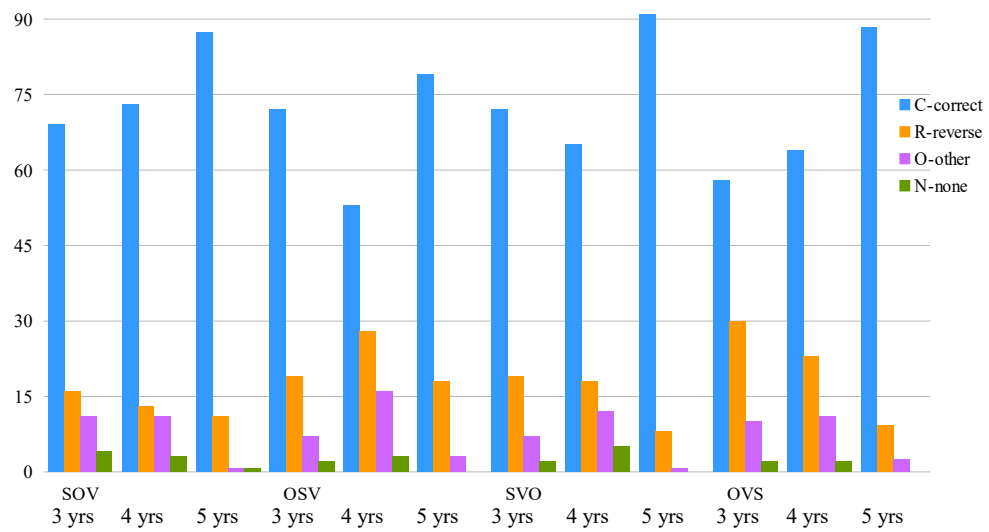
Figure 28: Results for Group III (mean age - 5;3years)



In Figure 28, it can be seen that the 5-year-old children made agent-patient reversal errors more often than other response errors. The percentages of correct answers were higher for this age group when compared to the younger age groups in the study.

When all the three graphs are compared (see Figure 29), it can be seen that while the 3-year-old children performed comparatively poorer in the OVS order, 4 and 5-year-old children had difficulty with the OSV order.

Figure 29: The percentage of all responses in all the word orders for all the age groups

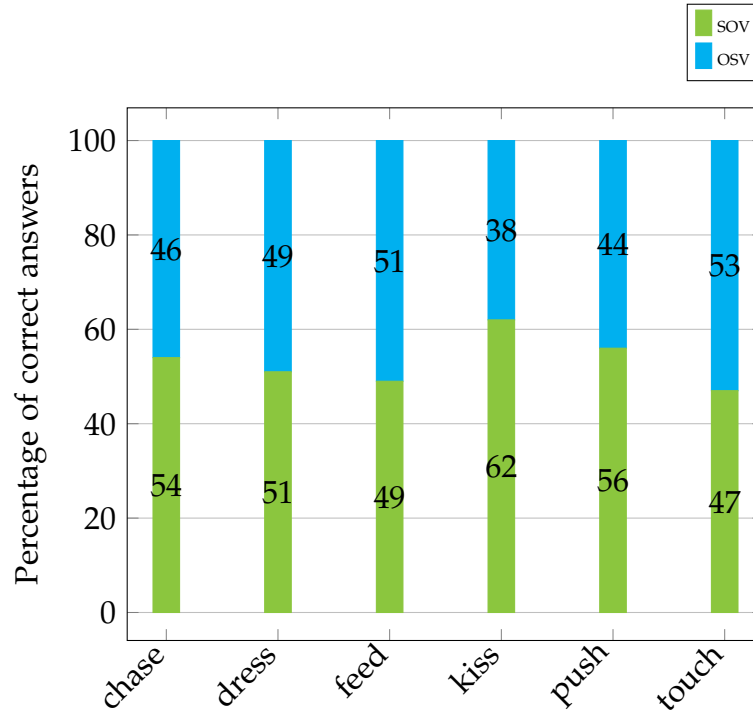


Since there were two groups of verbs, the performance with respect to the verbs and word order was compared. The percentage and the raw scores of correct answers for the verbs used with the SOV and the OSV order are given in Table 13 and in Figure 30.

Table 13: Percentage and raw score of the verbs used in SOV and OSV orders

	SOV		OSV	
Chase	38	54%	33	46%
Dress	43	51%	42	49%
Feed	50	49%	51	51%
Kiss	53	62%	33	38%
Push	46	56%	36	44%
Touch	45	47%	50	53%

Figure 30: Results for the verbs used in SOV and OSV orders

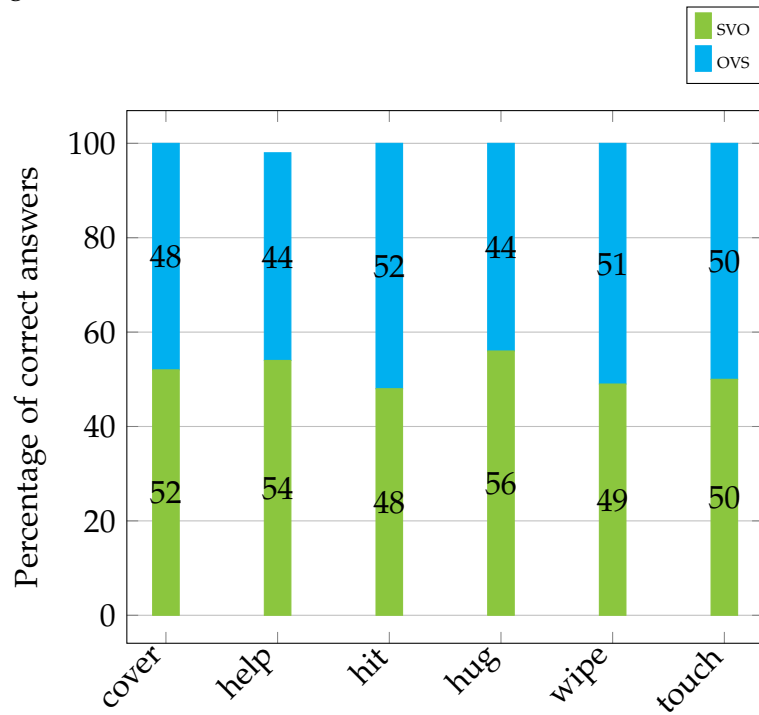


The percentage and raw scores of correct answers for the verbs in the SVO and in the OVS order are give in Table 14 and in Figure 31.

Table 14: Percentage and raw score of the verbs used in SVO and OVS orders

	SVO		OVS	
Cover	40	52%	37	48%
Help	43	54%	34	44%
Hit	44	48%	48	52%
Hug	52	56%	41	44%
Wipe	41	49%	43	51%
Touch	51	50%	50	50%

Figure 31: Results for the verbs used in SVO and OVS orders



No significant effect by the type of verb on word order was found.



### 5.3.3.2 Modeling the proportions of correct responses

Using the logistic regression model for bivariate data, the probability of the correct answers were calculated for each age group across each word order. It was found that the correct answers by the children in all the age groups were above chance for almost all word orders.

Group I children performed above chance for all the word orders, except for the performance of OVS order. There was no evidence for rejecting the at chance result for the OVS order since the confidence intervals were not above 0.5 (CI (95%) = (0.47, 0.68)) (see Table 15).

For the hypothesis on which the current experiment was made, results of Group I (mean age 3;5 years) is particularly relevant. From the overall results, since the scores were above chance, the null hypothesis, that the children by age 3 do not have knowledge of scrambling, was rejected across all the word orders tested. It can be concluded that native Malayalam children at the age of 3 has set the relevant parameters for clause internal scrambling.

Table 15: Probability of correct answers for Group I (mean age 3.5 years)

Order	Std Err	DF	Pr >  t	Mean	Low Mean	Upp Mean
OSV	0.234	1371	<.0001	0.7224	0.6217	0.8048
OVS	0.219	1371	0.1130	0.5861	0.4794	0.6853
SOV	0.230	1371	0.0003	0.6970	0.5943	0.7833
SVO	0.234	1371	<.0001	0.7224	0.6217	0.8048

Table 16 shows the results of Group II. The performance on all the orders except the OSV order were above chance. The same as the result of OVS in Group I, there was no evidence to reject chance performance for OSV because the confidence intervals were not above 0.5 (CI (95%) = (0.42, 0.63)).

Table 16: Probability of correct answers for Group II (mean age 4.3 years)

Order	Std Err	DF	Pr >  t	Mean	Low Mean	Upp Mean
OSV	0.217	1371	0.5343	0.5337	0.4277	0.6369
OVS	0.224	1371	0.0076	0.6454	0.5398	0.7385
SOV	0.236	1371	<.0001	0.7305	0.6304	0.8117
SVO	0.223	1371	0.012	0.6368	0.5309	0.7309

For Group III, performance for all sentence types was above chance (see Table 17).

Table 17: Probability of correct answers for Group III ( mean age 5.3 years)

Order	Std Err	DF	Pr >  t	Mean	Low Mean	Upp Mean
OSV	0.254	1371	<.0001	0.7968	0.7044	0.8658
OVS	0.308	1371	<.0001	0.8875	0.8116	0.9352
SOV	0.300	1371	<.0001	0.8793	0.8016	0.9292
SVO	0.338	1371	<.0001	0.9119	0.8421	0.9525

When the overall performance of correct answers with respect to age was analyzed, it was found that there was a significant difference between the performance of 3 and 5-year-old subjects ( $p < .0001$ ), and 4 and 5-year-olds ( $p < .0001$ ), but there was no statistically significant difference between the performance of 3 and 4-year-old subjects ( $p < .345$ ). The performance of the 5-year-old children was closer to the adult performance than that by the younger children.

#### 5.3.4 Discussion

The comprehension of scrambled sentences in the absence of a discourse context was investigated in Malayalam speaking children within the

age range of 3 and 5-years. The basic order SOV and the scrambled orders OSV, OVS and SVO were presented to children and adults. The comprehension was measured by the responses of a Picture Matching task.

The prediction based on the VEPS that children would be able to parse the scrambled sentences at the earliest testable age was borne out. The results revealed that children by the age of 3 have knowledge of scrambling. A gradual increase in performance by the older children when compared to the younger children was observed. Even though this did not affect the outcome of the predictions, a brief examination is due. The performance of the 5-year-old subjects was more similar to the adult performance than that by the 3-year-old subjects. One explanation for this could be that the younger children face difficulty in parsing the scrambled sentences in unnatural contexts (with no discourse). [Crain & Thornton \(2000\)](#) explain that if in an experimental task the syntax and pragmatics are incongruent then that counts as a flaw in the methodology. It was proposed that results from such methodology are not an apt measure that points to the syntactic abilities of child grammar. Linguistic stimuli are said to elicit accurate responses if provided within a natural discourse context. But it was seen that adults faced no difficulty in parsing scrambled sentences in the absence of discourse context. To perform an action in a novice environment, by parsing a sentence devoid of any context, one might need more cognitive skills in addition to syntactic knowledge. The performance by the adults provides an understanding that the adult grammar is making use of such extra-linguistic abilities. Since the 5-year-old children performed approximately like adults, it could be the case that the older children by the age of 5 have mastered the cognitive skills with which they can circumvent this pragmatic constraint and respond adequately. The 3-year-old children, on the other hand, performed in a way that reflects an immaturity in terms of parsing sentences in unnatural situations. This correlation would explain the statistically significant difference between the performance of 5-year-olds and 3-year-olds. The gradual increase in performance across age can thus be attributed to the maturation of the syntax-discourse interface. This conclusion can be compared to the results of [Schaeffer \(2000a\)](#). The 2-year-old Dutch chil-

dren had difficulty differentiating discourse related and non-discourse related referentiality and did not always scramble object DPs. Schaeffer proposed a delay in the development of pragmatics in the younger children in the study. It is crucial to remember that the average performance of the younger group in the Malayalam experiment was above chance for all the word orders, and at chance for OSV and OVS in 3-year-olds and in 4-year-olds respectively, despite the administration of the scrambled sentences without any discourse context. That being the case the comparatively poor performance by the 3-year-old children is comparable to the poor performance by the Dutch children in a situation that demands an advanced potential of the syntax-pragmatics interface. The 3-year-old Malayalam children displayed more rigidity than the adults. Expanding further, the above chance performance on most word orders should be taken as evidence for their knowledge on abstract syntactic constructions.

Ignoring the total results, for arguments sake if the comparatively lower performance by the young children is attributed to the fact that they have difficulty with scrambling, then the performance for the basic word order, SOV, should have been at ceiling. The fact that young children already know the basic word order of their native language (Brown, 1973; Sugisaki, 2008) is already proven. It was seen that the mean of the correct responses for the basic order, SOV, by the 3-year-old children was only 69% and not 100% like the adults. This means that the difficulty was due to the methodology and not due to syntax, and the gradual increase in performance in relation to age is due to the maturation in the extra-linguistic areas.

Another factor that demands explanation is the comparatively poor performance in some word orders. In one glance, both object-initial orders were troublesome for some children. Errors were more abundant in the OSV and the OVS orders for all the age groups, and the overriding error type was role-reversal. This indicates that in the OSV and the OVS orders some children interpreted the first NP as the agent. Fewer role-reversal errors in the SOV and in the SVO sentences confirm this interpretation.

According to Jayaseelan (1996) the canonical order, SOV, is derived via nested VP movements, from SVO to SOV. If that assumption is followed,

then the OVS and the OSV order are the results of a second order movement. The data show that children did not perform below chance levels for these orders which apparently involves more than one movement. This confirms the presence of scrambling in child grammar.

Since there was no discourse related information provided to the children it could be stated that even in the absence of pragmatic cues child grammar parses scrambled sentences effectively. The reason to say this is that many studies (Kidwai, 1999; Schaeffer, 2000a; Jayaseelan, 2008) have attributed scrambling to a feature checking movement, to [+Focus] or [+Referentiality]. The current experimental results of Malayalam scrambling are in line with the claim that the functioning of the syntax module matures independently of the changes in pragmatic development. The performance by Malayalam children was sufficiently good to disprove the overlapping effect of delayed pragmatics on syntax. If it is the under-developed syntax-pragmatics interface that resulted in the poor performance by the young children, then a different methodology where pragmatics is kept under check can isolate the performance of syntactic abilities in the scrambled sentences. That is, providing a discourse context and presenting the scrambled sentences in a natural environment would yield isolated results of syntactic performance alone. This conclusion has led to design the second experiment with a further developed methodology.

## 5.4 EXPERIMENT II

### 5.4.1 *Rationale*

Since the studies on Japanese scrambling (Otsu, 1994) showed that the presence of a discourse context is the key to better performance in scrambled sentences by young children, the current experiment was designed with a discourse context. The experimental design followed the suggestion by Crain & Thornton (2000) that linguistic input out of context would sound unnatural to children, and that might affect their performance. However, it should be noted that children at the age of three did show the knowledge of scrambling without any discourse context in the previous experiment. Here the acquisition of scrambling was tested in 2-year-old children using scrambled sentences which were administered with discourse contexts.

The updated hypothesis was that the children at the age of 2 would parse scrambled sentences, and the performance would be higher with a discourse context. In line with VEPS the prediction was that, the parameters that determine the clause-internal middle scrambling involving A' movement in Malayalam are set at 2-years of age.

### 5.4.2 *Experimental Design*

The experiment designed was an act-out task.

#### 5.4.2.1 *Subjects*

Twenty children, 13 females and 7 males, within the age range of 2;4–3;0 (mean age: 2;8) were tested. All children were monolingual Malayalam native speakers who were enrolled in the Wisdom Valley Kindergarten at Trivandrum, Kerala. No child was excluded from the study, and none of the children had any history of speech, language, hearing and cognitive impairment. A younger age (2-years) than the lowest age tested in the Experiment I was chosen as all the age groups performed above chance in Experiment I. No adults were tested, for the reason that 100% performance by the adults was noted in the previous experiment.

### 5.4.2.2 Materials

Eight toy animals: cow, dog, duck, hen, ox, cat, goat and horse were used. The toy animals were placed on a simulated platform of a farm, that is, a large plastic sheet with a picture of a farm printed on it was spread on a table.

### Test Sentences

Sixteen test sentences were constructed with four verbs. The simple past tense form was used, by adding the past tense marker *-i* and *-u* to the root form, for all the verbs in the test sentences. Root form of the verbs tested are given in Table 18.

Table 18: Verbs used in Experiment II

Verbs
<i>chavittuka</i> 'kick'
<i>thalluka</i> 'push'
<i>thadavuka</i> 'pet/massage'
<i>edukkuka</i> 'lift'

The test sentences were given in the basic order, SOV (129), and in three scrambled orders, OSV (130), SVO (131) and OVS (132).

- (129) *pashuvum thaaravum vellam kudikkuka aayirunnu.*  
 cow-CONJ duck-CONJ water drink-INF is-PRES-PROG.  
*appo pashu thaaravin-e thalli.*  
 then cow duck-ACC push-PAST  
 'Cow and duck were drinking water. Then the cow pushed the duck.'

- (130) *ee kozhiyum aadum koode nadannu*  
 this chicken-CONJ goat-CONJ together walk-INF  
*pokukayaayirunnu. appo aadin-e kozhi chavitti.*  
 go-PRES-PROG. Then goat-ACC chicken kick-PAST

'This goat and this chicken were walking together. Then the chicken kicked the goat.'

- (131) *kozhiyum thaaravum kalichondirunnappol veenu.*  
 chicken-CONJ duck-CONJ play-PROG-then fall-PAST.  
*appo thaaraavu eduthu kozhiy-e.*  
 Then duck lift-PAST chicken-ACC  
 'The duck and the chicken fell down while playing. Then the duck lifted the chicken.'

- (132) *kuthirayude aduthott chennu chiriche poocha. ennitt*  
 horse-GEN to near go-PAST smiled-ok cat.  
*kuthiray-e thadavi poocha*  
 Then horse-ACC massage-PAST cat-NOM  
 'The cat went near the horse and smiled, ok? Then the cat massaged the horse.'

#### 5.4.2.3 Procedure

The printed sheet of the farm was spread on a table, and all the toy animals were placed randomly in different places which simulated a scene of a farm with animals in it. Children were tested one by one, and they were made to sit facing the farm. The experimenter sat on the side and narrated the contexts with target sentences in it. The experimenter picked up the animals one by one to introduce them and later asked the child to give each animal by picking it up ('lift the cat and give it to me please'). After the introduction of characters, a warm up session was run in order for the child to understand the act-out task. Intransitive sentences were used to avoid any unintentional training being given to children about word order and agent and patient roles. In the warm up session, the toy cat and dog from the characters were placed inside a fence (see Figure 32).



Figure 32: The experimental setup



The situation was narrated first to the child, and the experimenter modeled one part of the story and asked the child to act out the last part. The scene proceeded as follows:

- (133) Experimenter: -*"The cat and the dog had a fight today, and the cat ran away, like this"*. [Experimenter acted out the cat running away towards the end of table. And the second part was narrated.]  
*"Seeing this the dog walked away to the other side, now show me that"*.  
 [Upon saying this if the child hesitated to act out she/he was prompted to pick the dog and act out the walking away.]

After the training session, the test trials were ran. For each sentence a context was given first, as exemplified in (129) - (132). The sentences were presented in random order. The children were asked to act out the sentences and the sentences were repeated upon the request. Positive reinforcements, as verbal praise or claps were given for all responses.

Each child received 4 tokens per word order and therefore a total of 16 sentences.

### *Coding*

Responses were coded as C for correct if children acted out the sentence with the right agent and patient, and R for reverse if they reversed the agent and patient. The experimenter herself noted down the responses on a notebook which was placed on her lap away from the subject's view.

### 5.4.3 *Results*

The descriptive analysis of the data was carried out. The results came out as self explanatory, that is, out of 320 total responses only 3 responses were incorrect (reverse). Table 19 gives the percentages and raw scores of total correct and incorrect answers.

Table 19: Raw score and percentage of correct and incorrect answers

	Correct	Incorrect
Percentage	99.06%	0.94%
Count	317	3

The percentage of total correct and incorrect answers plotted graphically is given in Figure 33.

Figure 33: Results

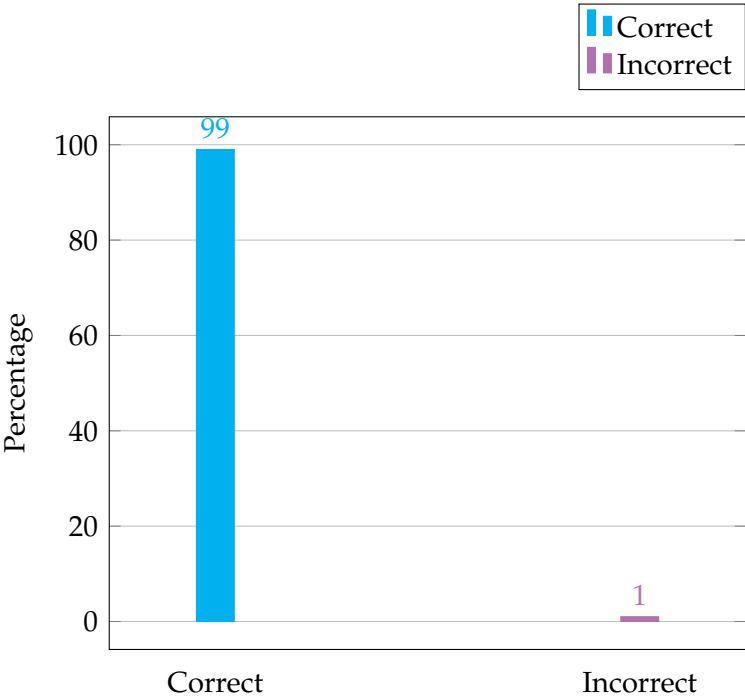


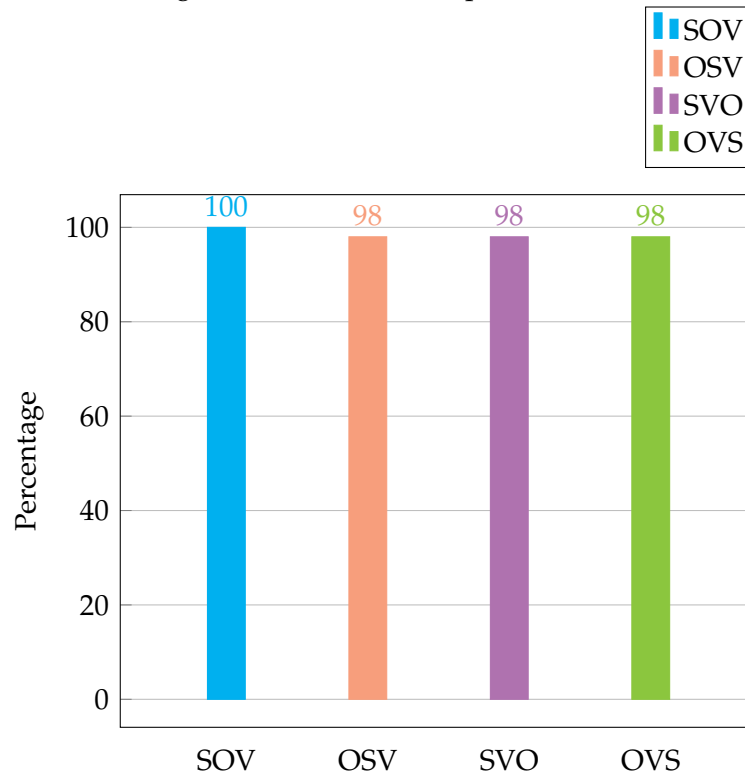
Table 20 gives the performance of children on each word order. The 2-year-old children performed 100% correctly for the SOV word order which is the canonical word order and 98.7% for all the scrambled word orders.

Table 20: Raw score and percentage of correct and incorrect answers per word order

	SOV	OSV	SVO	OVS
Correct percentage	100%	98.7%	98.7%	98.7%
Incorrect percentage	0%	1.25%	1.25%	1.25%
Correct raw scores	80	79	79	79
Incorrect raw scores	0	1	1	1

The performance of children on each order is plotted graphically in Figure 34.

Figure 34: Results of Experiment II



There was only one wrong response in each of the scrambled sentence types. This error was not produced by a single child but by three different children. All the three wrong responses were found for sentences with the verb *push*.

#### 5.4.4 Discussion

The comprehension of scrambled word orders was tested in 20 children within the age range of 2;4 - 3;0 years. The results showed that the children at age 2 comprehend scrambled sentences. The prediction based on VEPS that the 2-year-old children have set the parameters for scrambling was borne out. The second prediction that performance would

improve if scrambled sentences are provided with a discourse context was also borne out. The 2-year-old children performed like adults in parsing scrambled sentences. On comparing this result with the results of the previous experiment, it becomes clear that the poor performance by the older children in the previous experiment was due to the lack of a discourse context. The results tally with the fact that scrambling is the outcome of the various ways in which information structure is encoded in sentence structure. Since scrambling is most often described in association with information structure contexts, previous studies have attributed difficulties exhibited by children to the fact that their pragmatic abilities mature slowly. Any delay presented by young children in parsing scrambled sentences was explained using the under-development of the syntax-pragmatics interface. However, when children were tested in natural contexts, their performance reflected real competence with regard to pragmatics and syntax. These data can be used to argue that the ability to parse sentences and comprehend sentences using abstract syntactic rules of word order are achieved by children as young as 2-years of age.

A different possibility for the interpretation of scrambled sentences can be that the contexts for certain verbs could have given more information to children to identify the agent and the patient. For instance, the context for the verb *pet/message* was that the toy animal which was the patient was falling down or hurt, and the toy animal which was unhurt or standing up was carrying out the action (see Figure 35).

Figure 35: The context for the verb pet/massage



It was pointed out by Milada Walkova (personal communication) that in such a context the only possibility is that the animal which was fallen down or hurt would be the patient and children deduced agent and patient information from this cue. This possibility can be dismissed if the performance of verbs which did not have a contextual cue like this for the agent and patient are compared. For instance, in the contexts for verbs such as *kick* and *push* both the toy animals were in the same position. There was no difference in performance between these two conditions. The possibility of such an extra-linguistic cue did not affect the performance.

To conclude, my predictions were borne out, the children by the age of 2 did exhibit the knowledge of scrambling and the comprehension of scrambling increased when a discourse context was provided.

### 5.5 EXPERIMENT III

The previous two experiments provided conclusive evidence for the fact that the children who acquire Malayalam have the knowledge of scrambling by 2 years of age. The third experiment was carried out to investigate further if the child grammar contains one of the properties of scrambling, namely the reconstruction property. According to Saito

(1989), Japanese scrambling undergoes reconstruction. Parallel to the observation by Saito, it is observed that the anaphors and pronouns in Malayalam exhibit reconstruction in scrambled sentences (Jayaseelan, 1989), as I illustrate below.

The most common pronominal anaphor in Malayalam is *taan*, in singular form, and *tangal* in plural form. In one of the earlier accounts on anaphors in Malayalam, Mohanan (1982) identifies *taan* as a pronominal anaphor, with +anaphoric and +pronominal features. A few of its properties that Mohanan lists are: the anaphor must have an antecedent in the same sentence (134), but not in the same minimal NP or S (135), it must be c-commanded by the antecedent, and that the antecedent must be a subject (136).

- (134) *mohan<sub>i</sub> tante<sub>i</sub> bhaaryaye aaradhikkunnu*  
 mohan self-GEN wife-ACC worship-PRES-PROG  
 'Mohan<sub>i</sub> worships self's<sub>i</sub> wife.'

- (135) \**mohan tanne aaradhikkunnu*  
 mohan self-ACC worship-PRES-PROG  
 'Mohan worships self.'

- (136) *appu<sub>i</sub> tan-te<sub>i</sub> pusthakangal kumar sookshikkum ennu*  
 appu self-GEN book-PL kumar secure-FUT that  
*karuthi*  
 think-PAST  
 'Appu<sub>i</sub> thought Kumar will keep self's<sub>i</sub> books safe.'

*Taan* is considered to be a long distance anaphor according to Jayaseelan (1989) as it can take its antecedent outside the minimal clause. As exemplified in (137) the presence of any number of nominals between a potential antecedent and the anaphor does not block the binding relation of the anaphor to the subject antecedent.

- (137) [*mantri tan-te<sub>i</sub> bhaarya-ye nullunn-ath*] *senaanaayakan*  
 minister self-GEN wife-ACC pinch-Nominal army-chief  
*kandu ennu*] *raajaavu<sub>i</sub> vichaarichu*  
 saw COMP king thought  
 'The king<sub>i</sub> thought that the army-chief saw the minister pinching self's<sub>i</sub> wife.'

- example taken from Jayaseelan, 1989.

If *taan* occurs as the object NP, then it cannot co-refer to the subject NP in the same sentence (138) (Jayaseelan, 1989). In a sentence like (139), where *taan* itself is not an argument of the verb it can co-refer to the subject NP.

- (138) \**amma<sub>i</sub> tanne<sub>i</sub> nokki.*  
 mother self-ACC see-PAST  
 'Mother<sub>i</sub> looked at self<sub>i</sub>.'

- (139) *amma<sub>i</sub> tante<sub>i</sub> kunjine nokki.*  
 mother self-GEN baby-ACC see-PAST  
 'Mother<sub>i</sub> looked at her<sub>i</sub> baby.'

Jayaseelan makes another claim that *taan* is [+human]. However, both in written text and in spoken language instances where *taan* is referring to an animal are found. Some of these examples found from newspaper articles (www.mathrubhumi.com) are given in (140), (141) and (142).

- (140) *tante<sub>i</sub> shakthi thirichariyaathepokunna aana<sub>i</sub> ath*  
 self's power recognise-does-not elephant that  
*thirichariyunna anubhavamalle sharikkum madam pottal?*  
 recognise-does experience-is-not-Q really musth  
 'Isn't musth really the recognition of its<sub>i</sub> own powers by an elephant<sub>i</sub> who had not recognized it?'

- (141) *tante<sub>i</sub> aduthu nilkkumbol oraale polum*  
 self's near while-standing nobody mahout-ACC  
*paappaane thodaan anuvadikilla devi<sub>i</sub>.*  
 to-touch allow-not devi  
 'Devi<sub>i</sub> does not allow anyone to touch self's<sub>i</sub> mahout whenever (he) is near her.'

[Devi is the name of an elephant in this context.]

- (142) *tante<sub>i</sub> yajamaananu vendi pathinaaru kollam jeevichu,*  
 self-GEN owner-ACC for sixteen years lived,  
*oduviil addhehathinu vendi maricha oru naayayude<sub>i</sub> katha.*  
 at-last him-ACC for died a dog-DAT story.



'The story of a dog<sub>i</sub> who lived for sixteen years for self's<sub>i</sub> owner and who died for him.'

A general observation can be made about the scrambling of the anaphor *taan*, that is, as the anaphor is always adjoined to one of the arguments or adjuncts of the predicate, it can move only with that constituent. This can be illustrated with one of the test sentences that was used in the experiment. In the example (143) the anaphor *taan* modifies the PP, *tante veettil vech*.

- (143) *pashu<sub>i</sub> [tante<sub>i</sub> veettil vech] kuthiraye ummavechu.*  
 cow self-GEN home-LOC at horse-ACC kiss-PAST  
 'The cow<sub>i</sub> kissed the horse at self's<sub>i</sub> house.'

The only allowed orders of scrambling are the ones in which the anaphor moves as a part of a whole constituent along with the PP, as exemplified in (144)-(146).

- (144) *[tante<sub>i</sub> veettil vech] pashu<sub>i</sub> kuthiraye ummavechu.*  
 self-GEN home-LOC at cow horse-ACC kiss-PAST  
 'The cow<sub>i</sub> kissed the horse at self's<sub>i</sub> house.'

- (145) *kuthiraye [tante<sub>i</sub> veettil vech] pashu<sub>i</sub> ummavechu.*  
 horse-ACC self-GEN home-LOC at cow kiss-PAST  
 'The cow<sub>i</sub> kissed the horse at self's<sub>i</sub> house.'

- (146) *pashu<sub>i</sub> kuthiraye [tante<sub>i</sub> veettil vech] ummavechu.*  
 cow horse-ACC self-GEN home-LOC at kiss-PAST  
 'The cow<sub>i</sub> kissed the horse at self's<sub>i</sub> house.'

The intended meaning is lost if the anaphor is moved independently without the PP. In the example (147) *taan* modifies the horse and not the house. The sentence is syntactically well formed but with a different meaning. Notice the difference in meaning between (143) and (147) in the translation.

- (147) *\*pashu<sub>i</sub> tante<sub>i</sub> kuthiraye veettil vech ummavechu.*  
 cow self-GEN horse-ACC home-LOC at kiss-PAST  
 'The cow<sub>i</sub> kissed the self's<sub>i</sub> horse at house.'

This difference is crucial for the experiment, as the intended interpretation was that the children would pick the cow in (144) as the subject-antecedent of the anaphor within the PP, *self's house*. If the anaphor was moved independently, as in (147) then children could interpret the sentence as an SOV sentence and act out accordingly, i.e. they could show the cow kissing the horse.

In addition to losing the intended meaning, independent scrambling of the anaphor creates syntactically ill-formed sentences as well. When the anaphor modifies an object argument, this ungrammaticality is evident, as the following examples demonstrate. In the example (148) the anaphor *tante* modifies the object argument, unlike an adjunct PP as in the previous examples.

- (148) *amma<sub>i</sub> tante<sub>i</sub> kunjine eduthu.*  
 mother self-GEN baby-ACC pick-PAST  
 '(The) mother<sub>i</sub> picked up self's<sub>i</sub> baby.'

When the anaphor is split from the object argument it modifies, the sentence becomes ungrammatical (149). It has to be noted here that even when the subject antecedent, *amma*, is linearly present immediately preceding the anaphor, the sentence remains ungrammatical.

- (149) \**kunjine amma<sub>i</sub> tante<sub>i</sub> eduthu.*  
 baby-ACC mother self-GEN pick-PAST

The constraints on scrambling of the anaphor vary according to the case marker and its modifying constituent. For instance, when *taan* is accusative marked as *tanne*, the position of object argument and the anaphor can be interchanged without losing grammaticality. The orders object-anaphor and anaphor-object are grammatical in (150) and (151) as opposed to the order object-anaphor being ungrammatical in (149)

- (150) *appu tanne<sub>i</sub> kandathaayi amma<sub>i</sub> paranju.*  
 appu self-ACC see-Nominalizer-was mother say-PAST  
 '(The) mother<sub>i</sub> said that Appu saw self<sub>i</sub>.'

- (151) *tanne<sub>i</sub> appu kandathaayi amma<sub>i</sub> paranju.*  
 self-ACC appu see-Nominalizer-was mother say-PAST  
 '(The) mother<sub>i</sub> said that Appu saw self<sub>i</sub>.'

A few examples with different types of case markers on the anaphor (nominative marked ((152) - (157)), dative marked ((158) - (161))) are given below for a thorough understanding.

- (152) *sita<sub>i</sub> [taan<sub>i</sub> bahu sundariyaanu] ennu raavananodu paranju.*  
 sita self very beautiful-is that ravanan-SOC say-PAST  
 'Sita<sub>i</sub> told Ravanan that self<sub>i</sub> is very beautiful.'
- (153) *sita<sub>i</sub> [bahu sundariyaanu taan<sub>i</sub>] ennu raavananodu paranju.*  
 sita very beautiful-is self that ravanan-SOC say-PAST  
 'Sita<sub>i</sub> told Ravanan that self<sub>i</sub> is very beautiful.'
- (154) *[bahu sundariyaanu taan] ennu sita raavananodu paranju.*  
 very beautiful-is self that sita ravanan-SOC say-PAST  
 'Sita<sub>i</sub> told Ravanan that self<sub>i</sub> is very beautiful.'
- (155) *[taan bahu sundariyaanennu] sita raavananodu paranju.*  
 self very beautiful-is-that seetha raavanan-SOC say-PAST  
 'Sita<sub>i</sub> told Ravanan that self<sub>i</sub> is very beautiful.'
- (156) *\*taan<sub>i</sub> sita<sub>i</sub> bahu sundariyaanennu raavananodu paranju.*  
 self sita very beautiful-is-that raavanan-SOC say-PAST

The following examples demonstrate the grammatical and ungrammatical conditions when the anaphor is dative marked.

- (157) *\*bahu sundariyaanennu sita<sub>i</sub> taan<sub>i</sub> raavananodu paranju.*  
 very beautiful-is-that sita self raavanan-SOC say-PAST
- (158) *sita<sub>i</sub> [tanikk<sub>i</sub> raamanod ulla abhinivesham]*  
 sita self-DAT raman-SOC has-RP desire  
*baalishamaayi karuthi.*  
 silly-is thought  
 'Sita<sub>i</sub> thought that self's<sub>i</sub> desire for Raman is silly.'
- (159) *sita<sub>i</sub> [raamanodu tanikk ulla abhinivesham]*  
 sita raman-SOC self-DAT has-RP desire  
*baalishamaayi karuthi.*  
 silly-is thought  
 'Sita<sub>i</sub> thought that self's<sub>i</sub> desire for Raman is silly.'

(160) \**tanikk<sub>i</sub>*    *sita<sub>i</sub>*    *raamanodu*    *ulla*    *abhinivesham*    *baalishamaayi*  
          self-DAT    sita    raman-SOC    has-RP    desire    silly-is  
          *karuthi*  
          thought

(161) \**raamanodu*    *sita<sub>i</sub>*    *tanikk<sub>i</sub>*    *ulla*    *abhinivesham*    *baalishamaayi*  
          raman-SOC    sita    self-DAT    has-RP    desire    silly-is  
          *karuthi.*  
          thought

Further details of this are not dealt with in this thesis, as it is out of the scope of the topic of investigation. The details of the experiment are reported in the next section.

#### 5.5.1 *Rationale*

The current experiment was designed following the work by [Murasugi & Kawamura \(2005\)](#) on the acquisition of scrambling involving sentences with an anaphor. From their study, Murasugi and Kawamura found that 2-year-old Japanese children lack the lexical knowledge of anaphors and could not parse the SOV sentences with anaphors. On the other hand, 3-year-old Japanese children could parse both the SOV sentences, and the scrambled sentences which involved an anaphor. This information was incorporated into the prediction for the Malayalam experiment. The prediction was that, the children who can parse the SOV sentences that involve a pronominal anaphor would be able to parse scrambled OSV sentences with scrambled pronominal anaphors, i.e., if children can parse Subject-Anaphor-Object-V order, then they would be able to parse Object-Anaphor-Subject-Verb order.

#### 5.5.2 *Experimental Design*

The experiment designed was an act-out task.

### *Subjects*

Nine children in the age range of 2;3-3;4 years (mean age: 2;9 years) were tested. The children were enrolled in the Wisdom Valley Kindergarten, Trivandrum, Kerala, where the testing took place. All the subjects were native speakers of Malayalam, who had no family history of any language, cognitive or developmental impairment. Six children who were tested first were excluded from the study because the experimenter provided the test sentences only once even when the children asked for repetition. It was noted that the inadequate performance by these six children was because they did not get adequate exposure of the test sentences to parse them. The mistake became evident, as from the seventh child onward the performance improved when the test sentences were repeated. Only the children who were tested after correcting this error are included in the analysis.

### *Materials*

Six toy animal characters (horse, duck, chicken, cat, cow, and goat) were used. Pictures of houses, made out of paper, for individual animals, were propped up on a table.

### *Test Sentences*

The test sentences, both transitive and intransitive, were constructed using three different verbs. The list of verbs in their root form are given in Table 21. The simple past tense form of the verbs was used in the test sentences.

Table 21: Verbs used in Experiment III

Transitive	Intransitive
<i>thalluka</i> 'push'	<i>nadannupokuka</i> 'walk away'
<i>idikkuka</i> 'hit'	<i>odippokuka</i> 'run away'
<i>ummavekkuka</i> 'kiss'	<i>uranguka</i> sleep

The test sentences included: SOV sentences (162), intransitive sentences with the anaphor (163), SOV sentences with the anaphor (164), and scrambled OSV sentences with the scrambled anaphor (165). Each child received all the word orders with three different verbs for a total of 12 sentences. The sentences were presented in random order.

(162) *pashu kuthiray-e idichu.*  
 cow horse-ACC hit-PAST  
 'The cow hit the horse.'

(163) *poocha tant-e veettil kidannurangi.*  
 cat self-GEN house-LOC sleep-PAST  
 'The cat slept in self's house.'

(164) *pashu tante veettil vech kuthiray-e ummavechu.*  
 cow self-GEN home-LOC at horse-ACC kiss-PAST  
 'The cow<sub>i</sub> kissed the horse at self's<sub>i</sub> house.'

(165) *kozhiy-e tante<sub>i</sub> veettil vech thaaraav<sub>i</sub> ummavechu.*  
 hen-ACC self-GEN house-LOC at duck kiss-PAST  
 'The duck<sub>i</sub> kissed the hen at self's<sub>i</sub> house.'

### Procedure

On a table, the pictures of the houses were propped up. The toy animals were kept away from the table at the beginning of the experiment. For each sentence, the toy animals were lined up on one corner of the table, so that the children had to act out the sentence by picking up the animals and placing them near the respective houses. The target sentences were embedded in a script which was narrated to the children by the experimenter. After each target sentence, the children were instructed to act-out the sentence. The intransitive sentences in the script were first modeled by the experimenter. An example of one of the scripts is given below.

(166) [The cow and the horse toy animals were kept on one corner of the table, away from their respective houses.]

Experimenter: -"*pashuvum kuthirayum koottukaaraayirunne, avaru ayalkkaarum aayirunnu. pakshe chilappo okke avaru vazhakkidum,*

*ketto? ith pashoonte veedaane, ithu kuthirayude veedum.*" ('The cow and the horse are neighbors and they are friends, but sometimes they fight, ok? This is the cow's home and this is the horse's home.')

[The experimenter points to the respective picture props.]

*-“ippo kuthira tante veettilekk odippoyi.*" ('Now the horse ran to the self's home.')

[The experimenter modeled this action.]

*-“athu kandappo pashu tante veettilekk nadannu poyi (), athu kaaniche.*" ('And seeing that, the cow walked away to self's house, show me that.')

[The experimenter prompted the child to act out this sentence. The animals were put back into the place away from house props. The child was asked to point once more to show which is the horse's house and the cow's house. This was to make sure that the children knew which house was whose. And then the narration continued.]

*-“oru divasam pashuvum kuthirayum vazhakkittu. pashu kuthiraye idichu, ath kaaniche. shari.*" ('One day the cow and the horse fought. The cow hit the horse, show me that. OK.')

*-“avaru ennittu ivideyellaam odinadannu vazhakkittu, appo kuthiraye thante veettilevach pashu thalli, ath kaaniche. shari.*" ('Then they ran around the place and kept fighting, and the cow pushed the horse at self's house, show me that. OK.')

*-“kurachu kazhinjappo avaru pinneyum koottukoodi, appo pashu thante veettilevach kuthiraye ummavechu, ini ath kaaniche.*" ('After a while they became friends again, so the cow kissed the horse at self's house, show me that.')

All the experimental sentences which appear in (166) are glossed from (167) to (170).

- (167) *pashu tante veettilekk nadannupoyi.*  
 cow self-GEN home-to walk-PAST  
 'The cow walked to his house.'

- (168) *pashu kuthiray-e idichu.*  
 cow horse-ACC hit-PAST

'The cow hit the horse.'

- (169) *kuthiray-e tante veettilvech pashu thalli.*  
 horse-ACC self-GEN home-at cow push-PAST  
 'The cow<sub>i</sub> pushed the horse at his<sub>i</sub> house.'

- (170) *pashu tante veettilvech kuthiray-e ummavechu.*  
 cow self-GEN home-at horse-ACC kiss-PAST  
 'The cow<sub>i</sub> kissed the horse at his<sub>i</sub> house.'

Every time the experimenter prompted the child to act out the sentence, and the sentences were repeated on demand. The children were given positive feedback by the experimenter after each task, irrespective of the response being correct or incorrect.

### *Coding*

For the SOV sentences, responses were coded as correct if the right agent and patient were used in the act-out. The responses were coded as incorrect if the agent and patient were reversed. For the anaphor intransitive sentences if the act-out was carried out with the right agent and the respective house it was coded as correct, and if the wrong house was used it was coded as incorrect. For the anaphor SOV and the scrambled anaphor OSV sentences, responses were coded as correct when the children acted out the sentences using the right interpretation of the subject-antecedent, and if the wrong interpretation of the subject-antecedent was used for the act-out, then it was coded as incorrect. If the children acted out the sentences by ignoring the anaphor in these sentences that was also coded as incorrect.

### 5.5.3 *Results*

#### *Descriptive analysis*

Individual raw scores of performance across each word order are given in Table 22.



Table 22: Individual raw score per sentence type

Age in years	SOV	anaphor intransitive	anaphor SOV	anaphor OSV
2.3	3/3	3/3	2/3	2/3
2.4	2/3	3/3	2/3	2/3
2.8	3/3	3/3	3/3	3/3
2.9	3/3	3/3	2/3	2/3
3.0	2/3	3/3	2/3	2/3
3.0	3/3	3/3	2/3	2/3
3.2	3/3	3/3	1/3	3/3
3.3	2/3	3/3	2/3	2/3
3.4	2/3	3/3	3/3	2/3

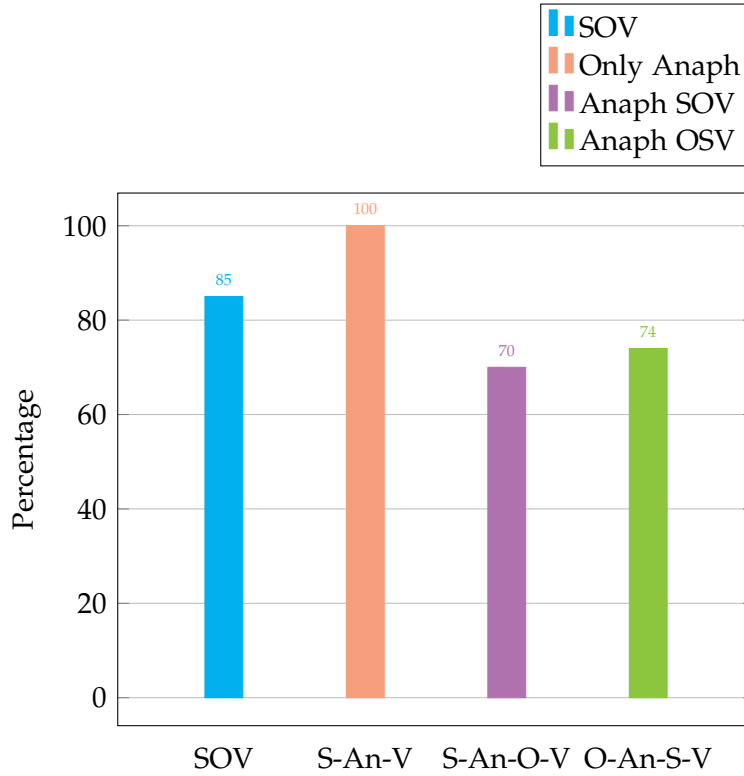
As it can be seen from the table, all the children performed at ceiling for the anaphor intransitive sentences. All the subjects parsed the anaphor sentences both in the basic word order and in the scrambled word order for more than 50% of the time except one child. The youngest subjects also parsed the agent and patient correctly along with the anaphor and its antecedent for more than 50%. The total raw scores and the percentage of correct answers per word order are given in Table 23.

Table 23: Raw score and percentage of correct answers

Word order	SOV	anaphor intransitive	anaphor SOV	anaphor OSV
Raw score	23/27	27/27	19/27	20/27
Percentage	85%	100%	70%	74%

The graph of the total raw score and the percentage of correct answers are given in Figure 36.

Figure 36: Results of Experiment III



#### *Modeling of proportion of correct answers*

All of the confidence intervals in the percentage of correct answers for each sentence type were above 50%. There was no difference between the performance of children in the sentences with an anaphor SOV, CI (95%) = (0.51,0.85) and in the sentences with a scrambled anaphor, CI (95%) = (0.54,0.87)), indicating that the 2-year-old children parsed the scrambled sentences with anaphor.

#### *5.5.4 Discussion*

Nine native Malayalam children within the age range of 2;3 - 3;4 years were tested for the comprehension of scrambled sentences including

anaphors. The results suggest that children by the age of 2 do have the knowledge of scrambling.

The experiment for Japanese by Murasugi & Kawamura (2005) had different results from the current experiment in Malayalam. The active sentence with the anaphor *zibun* was not parsed by the 2-year-old children in the Japanese experiment which led the authors to conclude that anaphors are not acquired until age 3. However, as it can be seen from the results of the Malayalam experiment, 2-year-old children did interpret the anaphor in both canonical and scrambled word order. The most probable factor for this discrepancy in results would be that no discourse context was provided by Murasugi and Kawamura in their experiment, whereas in the Malayalam experiment the target sentences were carefully embedded in a script with which children could gain adequate contextual information.

Murasugi and Kawamura explain the syntax of scrambled anaphor sentences using the reconstruction proposal from Saito (1989). Keeping aside that proposal, the results from the Malayalam anaphor scrambling experiment can be analyzed using the contrast that Chomsky (1995) proposed about the computing of linear order. Chomsky's proposal was that as the Copy Theory of displacement is an inevitable consequence of the Strong Minimalist Theory, and this aids in getting rid of operations like reconstruction. After Merge, once a syntactic object is displaced, it leaves a copy and the construction that reaches the Conceptual-Intentional (CI) interface would still be interpretable. Example (171) is one of the test sentences in the scrambled order, Obj-Anaphor-Sub-Verb, from the basic order, Sub-Anaphor-Obj-Verb.

- (171) *kozhiye      thante      veettil      vech      thaarav      ummavechu.*  
       hen-ACC self-GEN home-LOC at    duck      kissed  
       'The duck kissed the hen at self's house.'

The linear order provides an NP, *the hen*, immediately preceding the anaphor, and this would be the simplest computation for the anaphor to be interpreted if one goes by linear order. The structure dependent complex computation would require that the anaphor, *self*, be only interpreted with its subject antecedent, *the duck*. So the proposal is that at the CI interface a sentence with a displaced constituent is not parsed by

reconstruction, but by taking into account the copies of the displaced constituents and carrying out the structural computation. In the adult grammar this proposal holds true, the anaphor is interpreted with the remote NP, *the duck*. Hence, if the child grammar is adult-like in this respect then children should be able to parse sentences like (171). The evidence from the current experiment suggests that this is the case. Child grammar adheres to the structural computation rather than the linear order computation like adult grammar. This result supports the proposal by the Strong Minimalist Theory that linguistic rules are structure dependent.

The results point towards the fact that binding principles are already functioning in young children, and that they parse sentences abiding by these principles. The long distance anaphor *tante* has the feature of being subject oriented. When the children parsed the anaphor scrambled sentences, they showed the knowledge of the subject-antecedent feature of the anaphor. Even when the patient/object NP was linearly closer to the anaphor in the sentences, the children did not misinterpret it as the antecedent.

This study is the first, to my knowledge, that investigated the acquisition of anaphors and scrambling in Malayalam. The results of this experiment lead to the conclusions that Malayalam children have the lexical knowledge of the anaphor *taan*, and the knowledge of scrambling by the age of 2.

## 5.6 GENERAL DISCUSSION

### 5.6.1 Overview

Three experiments were conducted in native Malayalam children to investigate the acquisition of scrambling. In Experiment I the comprehension of scrambling without a discourse context was tested in 3, 4 and 5-year-old children. The subjects were given a picture matching task for linguistic stimuli which consisted of sentences with scrambled word orders (OSV, SVO, OVS) and the basic word order (SOV). The overall results showed that the youngest participants, the 3-year-old subjects, did comprehend all the sentences in most of the word orders. Their mean

performance was above chance level for all most word orders. While the linguistic stimuli in Experiment I were delivered without any discourse context, the sentences in Experiment II were provided with a context. The same word orders as Experiment I were used in an act-out task. The subjects for Experiment II were twenty 2-year-old native speakers of Malayalam. The results showed at ceiling performance by the 2-year-old children for all the word orders. Both these experiment results are conclusive of the fact that by 2-years of age the native Malayalam children do have the knowledge of scrambling.

The Japanese children were found to be acquiring the basic word order early and the scrambled word orders late (Hayashibe, 1975). Otsu (1994) contradicted this finding by using contextual cues, and likewise the results from the Experiment II proved that the acquisition of scrambling in Malayalam is early. In Experiment I, the overall performance was reduced slightly for the 4-year-old children, the reasons for which are unknown at this moment. Comparing the performance on each word order, the basic word order, SOV, received maximum correct responses and the order Object-Subject-Verb received the minimum number of correct responses. However, this was not statistically significant. There was also a gradual improvement in performance by the children with the advancement of age. The results from the Malayalam experiments are in agreement with the results from the Japanese experiments on the acquisition of scrambling by Sugisaki & Isobe (2001).

The experimental results provide empirical support to VEPS (Wexler, 1998). The children have set the various parameters for parsing scrambled sentences by the age of 2. The age of acquisition which is proposed as 2 is not the final definite claim, but it is only with respect to the current experimental results. It is to be investigated whether children also know the constraints of scrambling, whether they know that scrambling out of an embedded clause and infinitival clause are ungrammatical in Malayalam. That would be an added empirical evidence to the theories of language acquisition. The results from Experiment I and II suggest a maturation pattern in terms of the ability to parse non-canonical word orders without discourse context. Adults showed at ceiling performance for scrambled orders, but children from 3 years to 5 years of age showed

a gradual increase in performance, this is suggestive of the maturation of pragmatics.

The third experimental results revealed the knowledge of reconstruction involved in scrambled anaphor sentences. The children by the age of 2 exhibited the knowledge of this property and identified the antecedent of the anaphor correctly.

### 5.6.2 *Relativized Minimality and the Malayalam Child Grammar*

When the Relativized Minimality (RM) hypothesis is applied to Malayalam child grammar, some conflicting findings emerge. In order for the RM effect to apply first, it has to be made sure that scrambling in Malayalam is the result of movement. Following Jayaseelan (2008), Malayalam scrambling is the result of movement. Even though all the scrambled sentences used in the experiment are derived via movement only the order OSV qualifies for a minimality effect. The orders OVS and SVO were proposed to involve remnant VP movement in Jayaseelan's analysis and hence the possibility of intervening constituents does not hold (see Figures 8 and 9 for the derivation). Since Jayaseelan follows Kayne's analysis, every scrambled sentence then, in fact, includes two movements: first from the SVO to SOV and then further movement. So an OSV sentence is base-generated as an SVO sentence. After the first movement it becomes SOV, the basic word order, and then from there further movement takes place for the scrambled surface order, OSV. Similar is the case in OVS order. Even if the number of movements under Kayne's approach are not detailed here, the fact that the object moves across the subject in the OSV order stands valid. For the order SOV, the object is a potential intervener between the subject and its trace. If both arguments share the lexical feature +NP then, by the inclusion configuration analysis of RM, the child is expected to face difficulty in processing these sentences, and so they should be acquired late (Friedmann et al., 2009). The results from the experiments in Part IV of this thesis conducted on the acquisition of scrambling in Malayalam are contrary to these predictions. The children by the age of 2 did acquire all the scrambled word orders without showing any significant delay for any one word order type.

As a first possibility, it can be argued that since the subject in the SOV order is +Top and the object is +Foc, no intervention effect holds as in (172).

(172) +Top, +NP.....+Foc,+NP ..... <+Top, +NP>

However not all anti-symmetric analyses of Malayalam SOV (or SOV in any other language, in fact) assign a +Foc feature to the object. In that case, child RM should hold.

The second possibility, that child grammar is as efficient as the adult grammar in parsing scrambled sentences which involve constituents with similar features crossing over another, seems a more plausible explanation of the Malayalam data. Clearly adult grammar faces no difficulties with scrambled sentences. Regarding child grammar, the proposal that RM gives place to a delay in the acquisition of sentences with a potential intervener does not seem to hold for Malayalam. The results from all three experiments make it clear that 2 and 3-year-old children are capable of parsing scrambled sentences, both simple and complex ones.

## Part V

# CONCLUSION



## CONCLUSION AND FUTURE RESEARCH

---

The aim of the investigation reported in this thesis was to provide empirical evidence for the early acquisition of word order from two least explored languages. Experiments were conducted in Hindi-Urdu, an Indic language, and in Malayalam, a Dravidian language. The acquisition of the linear ordering of constituents was investigated in children from 19 months of age to 5 years old. In 19-month old infants, who were acquiring Hindi-Urdu, the acquisition of the OV order was tested using the preferential looking paradigm and the weird word order. The results showed that children by 19 months of age has set the parameter for basic OV order correctly. The cues that aid children to isolate OV order from an input that consists of canonical and non-canonical orders are unclear.

The results from the Hindi-Urdu experiment contradicted the claim by the usage-based theories (Tomasello, 1992) that word order is learned through rote memory from the lexical items of a verb. Young infants parsed grammatical sentences with pseudo-verbs and this refuted the claim by the lexical-learning hypothesis.

Children who were acquiring Malayalam were tested for their knowledge on clause internal scrambling. The results show that children by the age of 2 do have the knowledge of scrambling and the subject-antecedent relation of anaphors. The three experiments conducted in Malayalam can be used to analyze the performance versus competence hypotheses. Children even at the age of 5 were not able to perform according to their competence when the methodology was flawed. Test sentences were presented in an unnatural context without discourse information. However, when discourse contexts were provided, 2-year-old children were able to perform adequately using their competence.

The 5-year-old children performed better than the 3-year-olds even in the absence of discourse contexts. The adults performed at ceiling without discourse cues. So as Crain & Thornton (2000) suggested, the

performance of pragmatic skills are maturing and the child is competent at early ages with respect to grammar. Further investigation needs to be carried out about the acquisition of the intricate rules of scrambling in Malayalam. This thesis only investigated clause internal middle scrambling. Clause-internal short scrambling and further constraints on scrambling were not researched upon. How children perform in complex sentences like embedded clauses, relative clauses, etc. must be investigated further for a complete picture.

The results from the experiments reported in this thesis provide evidence for Very Early Parameter Setting (Wexler, 1998). The claim that the language faculty opts for structural computation rather than linear order computation (Chomsky, 1995, 2013, 2016) was found to be empirically true in child grammar when the results from the acquisition of scrambled anaphor sentences were considered.

The predictions of Relativized Minimality in child grammar (Friedmann et al., 2009) were critically analyzed using the results from the Malayalam experiments. Child grammar was found to be parsing sentences that qualify for minimality effect, i.e., sentences in the OSV and SOV order. Following the derivations of these orders by Jayaseelan (2001), it is apparent that according to the RM hypothesis sentences in these orders should have been difficult for children to parse. But the results from all three experiments in Malayalam showed otherwise. The results from the Malayalam Experiment I showed at chance performance for the OSV order by 3-year-old children, but the sentences were presented without discourse context. In Experiment II when discourse context was provided children performed at ceiling. This reveals that the difficulty with the sentences in the OSV order was due to the lack of discourse context of the scrambled order, and not due to minimality effect. These results were conflicting with the claim that child grammar would reject sentences with minimality effects.

#### FUTURE RESEARCH

The experiment conducted in French (Franck et al., 2013) provided evidence for the basic VO order setting in an SVO language which consists of no case markers and no scrambling. The next experiment, the one

which is part of this thesis, investigated the same in an SOV language with case markers and free word order. It would be noteworthy to investigate the performance of children who acquire an SVO language with case markers and relative freedom of word order. A third experiment is proposed to be carried out in Greek which is an SVO language with case markers. Results from this experiment would be beneficial to compare/contrast the infants' knowledge of VO order in an SVO language without case markers, French, and in an SVO language with case markers, Greek.

It would also be interesting to investigate performance at younger ages with the very same experiments.

To conclude, the findings in this thesis support the school of thought that language faculty is innately present in infants, and that the children build grammar at very early ages. The results from both Hindi-Urdu and Malayalam experiments are the first of its kind from these languages to the best of my knowledge.

Part VI

APPENDIX

## APPENDIX

## A.1 MALAYALAM EXPERIMENT I TEST SENTENCES

## SOV sentences

1. *aniyathi chechiye thallunnu.*  
younger sister elder sister-ACC push-PRES  
'The younger sister is pushing the elder sister.'
2. *aniyathi ammaye ummavekkunnu.*  
younger sister mother-ACC kiss-place-PRES  
'The younger sister is kissing the mother.'
3. *muthashi aniyathik vaarikodukkunnu.*  
grandmother younger sister-DAT feed-PRES  
'The grandmother is feeding the younger sister.'
4. *amma chechiye odikkunnu.*  
mother elder sister-ACC chase-PRES  
'The mother is chasing the elder sister.'
5. *amma muthashiye orukkunnu.*  
mother grandmother-ACC dress-PRES  
'The mother is dressing the grandmother.'
6. *aniyathi muthashiye thodunnu.*  
younger sister grandmother-ACC touch-PRES  
'The younger sister is touching the grandmother.'

## OSV sentences

1. *chechiye aniyathi thallunnu.*  
elder sister-ACC younger sister push-PRES

'The younger sister is pushing the elder sister.'

2. *ammaye aniyathi ummaṭṭekkunnu.*  
mother-ACC younger sister kiss-place-PRES

'The younger sister is kissing the mother.'

3. *aniyathik muthashi vaarikodukkunnu.*  
younger sister-DAT grandmother feed-PRES

'The grandmother is feeding the younger sister.'

4. *chechiye amma odikkunnu.*  
elder sister-ACC mother chase-PRES

'The mother is chasing the elder sister.'

5. *muthashiye amma orukkunnu.*  
grandmother-ACC mother dress-PRES

'The mother is dressing the grandmother.'

6. *muthashiye aniyathi thodunnu.*  
grandmother-ACC younger sister touch-PRES

'The younger sister is touching the grandmother.'

#### SVO sentences

1. *chettan thorthikodukkunnu achanu.*  
elder brother wipe-PRES father-DAT

'The elder brother is wiping the father.'

2. *achan puthappikkunnu aniyane.*  
father cover-PRES younger brother-ACC

'The father is covering the younger brother.'

3. *achan sahayikkunnu muthshane.*  
father help-PRES grandfather-ACC

'The father is helping the grandfather.'

4. *achan thodunnu aniyane.*  
father touch-PRES younger brother-ACC

'The father is touching the younger brother.'

5. *aniyan adikkunnu chettane.*  
 younger brother hit-PRES elder brother-ACC  
 'The younger brother is hitting the elder brother.'

*OVS sentences*

1. *aniyane kettipidikkunnu muthshan.*  
 younger brother-ACC hug-PRES grandfather  
 'The grand father is hugging the younger brother.'
2. *achanu thorthikodukkunnu chettan.*  
 father-DAT wipe-PRES elder brother  
 'The elder brother is wiping the father.'
3. *aniyane puthappikkunnu achan.*  
 younger brother-ACC cover-PRES father  
 'The father is covering the younger brother.'
4. *muthashane sahayikkunnu achan.*  
 grandfather-ACC help-PRES father  
 'The father is helping the grandfather.'
5. *aniyane thodunnu achan.*  
 younger brother-ACC touch-PRES father  
 'The father is touching the younger brother.'
6. *chettane adikkunnu aniyan.*  
 elder brother-ACC hit-PRES younger brother  
 'The younger brother is hitting the elder brother.'

A.2 MALAYALAM EXPERIMENT II TEST SENTENCES

*SOV sentences*

1. *poocha kozhiye chavitti.*  
 cat hen-ACC kick-PAST  
 'The cat kicked the hen.'

2. *kuthira thaaraavine thalli.*  
horse duck-ACC push-PAST  
'The horse pushed the duck.'
3. *kaala aadine eduthu.*  
ox goat-ACC lift-PAST  
'The ox lifted the goat.'
4. *kuthira kozhiye thadavi.*  
horse hen-ACC massage-PAST  
'The horse massaged the hen.'

*OSV sentences*

1. *aadine poocha chavitti.*  
goat-ACC cat kick-PAST  
'The cat kicked the goat.'
2. *naayaye thaaraavu thalli.*  
dog-ACC duck push-PAST  
'The duck pushed the dog.'
3. *poochaye naaya eduthu.*  
cat-ACC dog lift-PAST  
'The dog lifted the cat.'
4. *pattiye poocha thadavi.*  
dog-ACC cat massage-PAST  
'The cat massaged the dog.'

*SVO sentences*

1. *naaya chavitti aadine.*  
dog kick-PAST goat-ACC  
'The dog kicked the goat'.
2. *kuthira thalli aadine*  
horse push-PAST goat-ACC



'The horse pushed the goat.'

3. *thaaraavu eduthu kozhiye.*  
duck lift-PAST hen-ACC  
'The duck lifted the hen.'

4. *kaala thadavi thaaraavine.*  
ox message-PAST duck-ACC  
'The ox massaged the duck.'

#### OVS sentences

1. *naayaye chavitti kozhi.*  
dog-ACC kick-PAST hen  
'The hen kicked the dog.'
2. *pashuvine thalli aadu.*  
cow-ACC push-PAST goat  
'The goat pushed the cow.'
3. *kozhiye eduthu pashu.*  
hen-ACC lift-PAST cow  
'The cow lifted the hen.'
4. *kuthiraye thadavi poocha.*  
horse-ACC message-PAST cat  
'The cat massaged the horse.'

#### A.3 MALAYALAM EXPERIMENT III TEST SENTENCES

#### SOV sentences

1. *pashu kuthiraye idichu.*  
cow horse-ACC hit-PAST  
'The cow hit the horse.'
2. *poocha aadine thalli.*  
cat goat-ACC push-PAST

'The cat pushed the goat.'

3. *aadu poochaye ummavechu.*  
goat cat-ACC kiss-PAST  
'The goat kissed the cat.'

*Anaphor SOV sentences*

1. *kozhi thante veetilvech thaaraavine thalli.*  
hen self-GEN home-at duck-ACC push-PAST  
'The hen<sub>i</sub> pushed the duck at his<sub>i</sub> house.'
2. *pashu thante veetilvech kuthiraye ummavechu.*  
cow self-GEN home-at horse-ACC kiss-PAST  
'The cow<sub>i</sub> kissed the horse at his<sub>i</sub> house.'
3. *thaaraavu thante veetilvech kozhiye idichu.*  
duck self-GEN home-at hen-ACC hit-PAST  
'The duck<sub>i</sub> hit the hen at his<sub>i</sub> house.'

*Anaphor OSV (scrambled) sentences*

1. *aadine thante veetilvech poocha idichu.*  
goat-ACC self-GEN home-at cat hit-PAST  
'The cat<sub>i</sub> hit the goat at his<sub>i</sub> house.'
2. *kuthiraye thante veetilvech pashu thalli.*  
horse-ACC self-GEN home-at cow push-PAST  
'The cow<sub>i</sub> pushed the horse at his<sub>i</sub> house.'
3. *kozhiye thante veetilvech thaaraav ummavechu.*  
hen-ACC self-GEN home-at duck kiss-PAST  
'The duck<sub>i</sub> kissed the hen at his<sub>i</sub> house.'

*Only Anaphor intransitive sentences*

1. *thaaraavu thante veedinuchuttum odinadannu.*  
duck self-GEN house-around ran-around

'The duck ran around his house.'

2. *poocha thante veettil kidannurangi.*  
 cat self-GEN home-at sleep-PAST

'The cat slept at his house.'

3. *pashu thante veettilekk nadannupoyi.*  
 cow self-GEN home-to walk-PAST

'The cow walked to his house.'

Part VII

BIBLIOGRAPHY

## BIBLIOGRAPHY

---

- Abbot Smith, K., Lieven, E., & Tomasello, M. (2001). What preschool children do and do not do with ungrammatical word orders. *Cognitive Development*, 16(2), 679–692.
- Abels, K. (2015). Word order. In A. Alexiadou & T. Kiss (Eds.), *Syntax - Theory and Analysis: An International Handbook* (pp. 1400–1448). Berlin: De Gruyter Mouton.
- Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language*, 26(02), 339–356.
- Akhtar, N. & Tomasello, M. (1997). Young childrens productivity with word order and verb morphology. *Developmental Psychology*, 33(6), 952.
- Allen, S. (1996). *Aspects of argument structure acquisition in Inuktitut*. Amsterdam: John Benjamins Publishing.
- Asher, R. E. & Kumari, T. (1997). *Malayalam*. London: Routledge.
- Avrutin, S. & Brun, D. (2001). The expression of specificity in a language without determiners: Evidence from Child Russian. In D. Anna H.-J., L. Dominguez, & A. Johansen (Eds.), *Proceedings of BUCLD*, volume 25 (pp. 70–81). Somerville, MA: Cascadilla Press.
- Baker, M. C. (2002). *The Atoms of Language: The Minds Hidden Rules of Grammar*. New York: Basic Books.
- Batman-Ratyosyan, N. & Stromswold, K. (1999). What Turkish acquisition tells us about underlying word order and scrambling. *University of Pennsylvania Working Papers in Linguistics*, 6(1), 37–54.
- Behrens, H. (1998). Where does the information go. In *MPI workshop on argument structure*. Nijmegen.

- Belletti, A. & Rizzi, L. (2013). Intervention in grammar and processing. In I. Caponigro & C. Cecchetto (Eds.), *From Grammar to Meaning: The Spontaneous Logicality of Language* (pp. 293–311). Cambridge: Cambridge University Press.
- Berman, R. A. & Armon-Lotem, S. (1995). How grammatical are early verbs? In *Annales Littéraires-Universite de Besancon*, volume 631 (pp. 17–60).: Dialogues d'Histoire Ancienne.
- Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the Development of Language* (pp. 279–362). New York: ERIC.
- Biberauer, T., Holmberg, A., Roberts, I., & Sheehan, M. (2009). *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.
- Biberauer, T. & Roberts, I. (2012). Towards a parameter hierarchy for auxiliaries: Diachronic considerations. *Cambridge Occasional Papers in Linguistics*, 6(9), 267–294.
- Borer, H. & Wexler, K. (1987). The Maturation Of Syntax. In T. Roeper & E. Williams (Eds.), *Parameter Setting* (pp. 123–172). Dordrecht: Reidel.
- Borer, H. & Wexler, K. (1992). Bi-unique relations and the maturation of grammatical principles. *Natural Language & Linguistic Theory*, 10(2), 147–189.
- Bošković, Ž. & Takahashi, D. (1998). Scrambling and last resort. *Linguistic Inquiry*, 29(3), 347–366.
- Brooks, P. J. & Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. *Developmental Psychology*, 35(1), 29.
- Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Berlin: Walter de Gruyter.

- Chomsky, N. (1993). On the nature, use and acquisition of language. In A. Clark & J. Toribio (Eds.), *Language and Meaning in Cognitive Science: Cognitive Issues and Semantic Theory*. Cambridge, MA: MIT Press.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (1999). Derivation by phase. In M. J. Kenstowicz (Ed.), *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press.
- Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1), 1–22.
- Chomsky, N. (2013). Problems of projection. *Lingua*, 130, 33–49.
- Chomsky, N. (2016). Minimal computation and the architecture of language. *Chinese Semiotic Studies*, 12(1), 13–24.
- Christophe, A., Guasti, M. T., & Nespor, M. (1997). Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, 12(5–6), 585–612.
- Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech*, 51(1–2), 61–75.
- Christophe, A., Nespor, M., Guasti, M. T., & Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science*, 6(2), 211–220.
- Clahsen, H. (1992). Learnability theory and the problem of development in language acquisition. In J. Weissborn, H. Goodluck, & T. Roeper (Eds.), *Theoretical Issues in Language Acquisition: Continuity and Change in Development* (pp. 53–76). Hillsdale, NJ: Lawrence Erlbaum.
- Crain, S. & Fodor, J. D. (1987). Sentence matching and overgeneration. *Cognition*, 26(2), 123–169.
- Crain, S. & Pietroski, P. (2001). Nature, nurture and Universal Grammar. *Linguistics and Philosophy*, 24(2), 139–186.

- Crain, S. & Thornton, R. (2000). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, MA: MIT Press.
- Culicover, P. W. (1997). *Principles and Parameters: An Introduction to Syntactic Theory*. Oxford: Oxford University Press.
- De Villiers, J. & Roeper, T. (2011). Introduction. In J. De Villiers & T. Roeper (Eds.), *Handbook of Generative Approaches to Language Acquisition* (pp. 1–12). Dordrecht: Springer.
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, 298(5600), 2013–2015.
- Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). Young German children's early syntactic competence: A preferential looking study. *Developmental Science*, 11(4), 575–582.
- Dodson, K. & Tomasello, M. (1998). Acquiring the transitive construction in English: The role of animacy and pronouns. *Journal of Child Language*, 25(03), 605–622.
- Dryer, M. S. & Haspelmath, M., Eds. (2013). *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Dyakonova, M. (2005). Information structure development: Evidence from the acquisition of word order in Russian and English. *Nordlyd*, 32(1).
- Felix, S. W. (1992). Language acquisition as a maturational process. In J. Weissborn, H. Goodluck, & T. Roeper (Eds.), *Theoretical Issues in Language Acquisition: Continuity and Change in Development* (pp. 25–51). Hillsdale, NJ: Lawrence Erlbaum.
- Féry, C. (2010). Indian languages as intonational 'phrase languages'. In I. Hasnain & S. Chaudhury (Eds.), *Festschrift to Honour Ramakant Agnihotri*. Delhi: Aakar Publisher.



- Fisher, C. (2002). Structural limits on verb mapping: the role of abstract structure in 2.5-year-olds interpretations of novel verbs. *Developmental Science*, 5(1), 55–64.
- Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(2), 143–149.
- Fisher, C., Hall, D. G., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, 92, 333–375.
- Fisher, C. & Tokura, H. (1996). Prosody in speech to infants: Direct and indirect acoustic cues to syntactic structure. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp. 343–363). Mahwah, NJ: Erlbaum.
- Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Franck, J., Millotte, S., & Lassotta, R. (2011). Early word order representations: Novel arguments against old contradictions. *Language Acquisition*, 18(2), 121–135.
- Franck, J., Millotte, S., Posada, A., & Rizzi, L. (2013). Abstract knowledge of word order by 19 months: An eye-tracking study. *Applied Psycholinguistics*, 34(02), 323–336.
- Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, 119(1), 67–88.
- Fukui, N. & Sakai, H. (2003). The visibility guideline for functional categories: Verb raising in Japanese and related issues. *Lingua*, 113(4), 321–375.
- Gathercole, V. C. M., Sebastián, E., & Soto, P. (1999). The early acquisition of Spanish verbal morphology: Across-the-board or piecemeal knowledge? *International Journal of Bilingualism*, 3(2-3), 133–182.
- Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules abstract knowledge of word order in early sentence comprehension. *Psychological Science*, 17(8), 684–691.

- Gervain, J., Nespor, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognitive Psychology*, 57(1), 56–74.
- Gianollo, C., Guardiano, C., & Longobardi, G. (2008). Three fundamental issues in parametric linguistics. In T. Biberauer (Ed.), *The Limits of Syntactic Variation* (pp. 109–142). Amsterdam: John Benjamins.
- Gibson, E. & Wexler, K. (1994). Triggers. *Linguistic Inquiry*, 25(3), 407–454.
- Göksun, T., Küntay, A. C., & Naigles, L. R. (2008). Turkish children use morphosyntactic bootstrapping in interpreting verb meaning. *Journal of Child Language*, 35(02), 291–323.
- Grebenyova, L. (2012). *Syntax, Semantics and Acquisition of Multiple Interrogatives: Who Wants What?* Amsterdam: John Benjamins.
- Greenberg, J. H. (1963). *Universals of Language*. Cambridge, MA: MIT Press.
- Grewendorf, G. & Sabel, J. (1999). Scrambling in German and Japanese: Adjunction versus multiple specifiers. *Natural Language & Linguistic Theory*, 17(1), 1–65.
- Guasti, M. T., Christophe, A., van Ooyen, B., & Nespor, M. (2001). Prelexical setting of the head-complement parameter through prosody. *Language Acquisition and Language Disorders*, 23, 231–248.
- Haider, H. (2000). Scrambling - what is the state of the art. In S. M. Powers & C. Hamann (Eds.), *The acquisition of Scrambling and Cliticization* (pp. 19–40). Dordrecht: Kluwer.
- Hakuta, K. (1977). Word order and particles in the acquisition of Japanese. *Papers and Reports on Child Language Development*, 13, 110–117.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579.

- Hayashibe, H. (1975). Word order and particles: A developmental study in Japanese. *Descriptive and Applied Linguistics*, 8(1), 18.
- Hirsh-Pasek, K. & Golinkoff, R. M. (1999). *The Origins of Grammar: Evidence from Early Language Comprehension*. Cambridge, MA: MIT Press.
- Hu, S., Gavarró, A., Vernice, M., & Guasti, M. T. (2016). The acquisition of Chinese relative clauses: contrasting two theoretical approaches. *Journal of Child Language*, 43(01), 1–21.
- Ilic, T. & Deen, K. U. (2004). Object raising and cliticization in Serbo-Croatian child language. In J. van Kampen & S. Baauw (Eds.), *Proceedings of 5th Generative Approaches to Language Acquisition* (pp. 235–243). Utrecht: LOT.
- Ingham, R. (1993). Critical influences on the acquisition of verb transitivity. In G. J. Turner & D. J. Messer (Eds.), *Critical influences on child language acquisition and development* (pp. 121–137). London: Maximilian Press.
- Isobe, M. (2009). Acquisition of head-internal relative clauses revisited. In J. Crawford, K. Otaki, & M. Takahashi (Eds.), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)* (pp. 88–97). Somerville, MA: Cascadilla Proceedings Project.
- Iwatate, S. (1981). Word order in early utterances of Japanese children. *Japanese Journal of Educational Psychology*, 29(2), 105–111.
- Jayaseelan, K. (1996). Question-word movement to focus and scrambling in Malayalam. *Linguistic Analysis*, 26, 63–83.
- Jayaseelan, K. A. (1989). Malayalam anaphors and the binding principles. In *Parametric studies in Malayalam Syntax*. New Delhi, Mumbai: Allied Publishers Limited.
- Jayaseelan, K. A. (2001). IP-internal topic and focus phrases. *Studia Linguistica*, 55(1), 39–75.

- Jayaseelan, K. A. (2008). Topic, focus and adverb positions in clause structure. *Nanzan Linguistics*, 4, 43–68.
- Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants preference for the predominant stress patterns of English words. *Child Development*, 64(3), 675–687.
- Kachru, Y. (1966). *An Introduction to Hindi Syntax*. Urbana, IL: Department of Linguistics, University of Illinois.
- Kachru, Y. (2006). *Hindi*. Amsterdam: John Benjamins Publishing.
- Kayne, R. S. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Kidwai, A. (1999). Word order and focus positions in universal grammar. In G. Rebuschi & L. Tuller (Eds.), *The Grammar of Focus* (pp. 213–244). Amsterdam: John Benjamins Publishing.
- Landau, B. & Gleitman, L. (1985). *Language and Experience*, volume 174. Cambridge, MA: Harvard University Press.
- Lee, J. N. & Naigles, L. R. (2008). Mandarin learners use syntactic bootstrapping in verb acquisition. *Cognition*, 106(2), 1028–1037.
- Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: experimental evidence for syntactic structure at 18 months. *Cognition*, 89(3), 295–303.
- Lightfoot, D. (2006). *How New Languages Emerge*. Cambridge: Cambridge University Press.
- Lust, B. & Chien, Y.-C. (1984). The structure of coordination in first language acquisition of Mandarin Chinese: Evidence for a universal. *Cognition*, 17(1), 49–83.
- Mahajan, A. K. (1990). *The A/A-bar distinction and movement theory*. PhD thesis, Massachusetts Institute of Technology.
- Manetta, E. (2012). Reconsidering rightward scrambling: Postverbal constituents in Hindi-Urdu. *Linguistic Inquiry*, 43(1), 43–74.

- Maratsos, M., Gudeman, R., Gerard-Ngo, P., & DeHart, G. (1987). A study in novel word learning: The productivity of the causative. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition* (pp. 89–114). Hillsdale, NJ: Erlbaum.
- Matsuo, A., Kita, S., Shinya, Y., Wood, G. C., & Naigles, L. (2012). Japanese two-year-olds use morphosyntax to learn novel verb meanings. *Journal of Child Language*, 39(03), 637–663.
- Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2007). French childrens use and correction of weird word orders: A constructivist account. *Journal of Child Language*, 34(02), 381–409.
- Mazuka, R. (1996). Can a grammatical parameter be set before the first word? prosodic contributions to early setting of a grammatical parameter. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp. 313–330). Mahwah, NJ: Erlbaum.
- Mohanan, K. P. (1982). Grammatical relations and clause structure in Malayalam. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations* (pp. 504–589). Cambridge, MA: MIT Press.
- Mohanan, K. P. & Mohanan, T. (1994). Issues in word order in South Asian languages: Enriched phrase structure or multidimensionality. In M. Butt, T. King Holloway, & G. Ramchand (Eds.), *Theoretical Perspectives on Word Order in South Asian Languages* (pp. 153–184). Stanford: Center for the Study of Language (CSLI).
- Mohanan, T. (1994). *Argument Structure in Hindi*. Stanford: Center for the Study of Language (CSLI).
- Morgan, J. L. & Demuth, K. (1996). *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Erlbaum.
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, 19(4), 498–550.

- Murasugi, K. & Kawamura, T. (2005). On the acquisition of scrambling in Japanese. In J. Sabel & M. Saito (Eds.), *The Free Word Order Phenomenon: Its Syntactic Sources and Diversity* (pp. 221–242). Berlin: Walter de Gruyter.
- Mykhaylyk, R. (2012). Factors contributing to child scrambling: evidence from Ukrainian. *Journal of Child Language*, 39(03), 553–579.
- Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17(02), 357–374.
- Narasimhan, B. & Dimroth, C. (2008). Word order and information status in child language. *Cognition*, 107(1), 317–329.
- O Grady, W. (2010). Emergentism. In P. C. Hogan (Ed.), *The Cambridge Encyclopedia of the Language Sciences* (pp. 274–76). Cambridge: Cambridge University Press.
- Olguin, R. & Tomasello, M. (1993). Twenty-five-month-old children do not have a grammatical category of verb. *Cognitive Development*, 8(3), 245–272.
- Omaki, A., Lassotta, R., Rizzi, L., & Franck, J. (2012). Input effects on parser development: Evidence from Japanese word order development. In *AMLaP 2012 Conference, September*.
- Otsu, Y. (1994). Early acquisition of scrambling in Japanese. In T. Hoekstra & B. D. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar* (pp. 253–264). Amsterdam: John Benjamins Publishing.
- Pierce, A. E. (2012). *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars*. Dordrecht: Kluwer.
- Pinker, S. (1984). *Language Learnability and Language Learning*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition* (pp. 399–442). Hillsdale, NJ: Erlbaum.

- Pinker, S. (1994). How could a child use verb syntax to learn verb semantics? *Lingua*, 92, 377–410.
- Pinker, S. (1995). *The Language Instinct: The New Science of Language and Mind*. London: Penguin.
- Pizzuto, E. & Caselli, M. C. (1994). The acquisition of Italian verb morphology in a cross-linguistic perspective. In Y. Levy (Ed.), *Other children, Other Languages: Issues in the Theory of Language Acquisition* (pp. 137–187). Hillsdale, NJ: Erlbaum.
- Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.
- Rizzi, L. (1993). Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition*, 3(4), 371–393.
- Roberts, I. (2012). Macroparameters and minimalism. In C. Galves, S. Cyrino, R. Lopez, F. Sandalo, & J. Avelar (Eds.), *Parameter Theory and Linguistic Change*. Oxford: Oxford University Press.
- Ross, J. R. (1967). *Constraints on variables in syntax*. PhD thesis, Massachusetts Institute of Technology.
- Rubino, R. B. & Pine, J. M. (1998). Subject-verb agreement in Brazilian Portuguese: what low error rates hide. *Journal of Child Language*, 25(1), 35–59.
- Saito, M. (1989). Scrambling as semantically vacuous A-bar movement. In M. R. Baltin & A. S. Kroch (Eds.), *Alternative Conceptions of Phrase Structure* (pp. 182–200). Chicago: The University of Chicago Press.
- Sano, K. (1977). An experimental study on the acquisition of Japanese simple sentences and cleft sentences. *Descriptive and Applied Linguistics*, 10, 213–233.
- Schaeffer, J. (2000a). Object scrambling and specificity in Dutch child language. In S. M. Powers & C. Hamann (Eds.), *The Acquisition of Scrambling and Cliticization* (pp. 71–93). Dordrecht: Kluwer.
- Schaeffer, J. C. (2000b). *The Acquisition of Direct Object Scrambling and Clitic Placement: Syntax and Pragmatics*. Amsterdam: John Benjamins.

- Schmitt, C. & Miller, K. (2010). Using comprehension methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental Methods in Language Acquisition Research* (pp. 35–56). Amsterdam: John Benjamins.
- Serrat, E. (1997). *Acquisition of verb category in Catalan*. PhD thesis, unpublished Thesis.
- Sheehan, M. (2014). Towards a parameter hierarchy for alignment. In R. E. Santana-LaBarge (Ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics (WCCFL 31)* (pp. 399–408). Somerville, MA: Cascadilla Press.
- Stoll, S. (1998). The role of aktionsart in the acquisition of Russian aspect. *First Language*, 18(54), 351–376.
- Stowell, T. A. (1981). *Origins of phrase structure*. PhD thesis, Massachusetts Institute of Technology.
- Sugisaki, K. (2008). Early acquisition of basic word order in Japanese. *Language Acquisition*, 15(3), 183–191.
- Sugisaki, K. & Isobe, M. (2001). What can child Japanese tell us about the syntax of scrambling? In K. Megerdooomian & L. A. Bar-el (Eds.), *Proceedings of the 20th West Coast Conference on Formal Linguistics (WCCFL 20)* (pp. 538–551). Somerville, MA: Cascadilla Press.
- Sugisaki, K. & Otsu, Y. (2011). Universal grammar and the acquisition of Japanese syntax. In J. De Villiers & T. Roeper (Eds.), *Handbook of Generative Approaches to Language Acquisition* (pp. 291–317). Dordrecht, NY: Springer.
- Tada, H. (1993). *A/A-bar partition in derivation*. PhD thesis, Massachusetts Institute of Technology.
- Tomasello, M. (1992). *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74(3), 209–253.



- Tomasello, M., Akhtar, N., Dodson, K., & Rekau, L. (1997). Differential productivity in young childrens use of nouns and verbs. *Journal of Child Language*, 24(02), 373–387.
- Tomasello, M. & Brooks, P. J. (1999). Early syntactic development: A construction grammar approach. In M. Barrett (Ed.), *The Development of Language*. London and New York: Psychology Press.
- Varma, A. R. R. (2000). *Kerala Paniniyam*. Kottayam: DC Books.
- Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, 106(1), 23–79.