

2475/4068

MONASH UNIVERSITY
THESIS ACCEPTED IN SATISFACTION OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ON..... 22 July 2003

.....
Sec. Research Graduate School Committee

Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing for the purposes of research, criticism or review. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.

Errata

p. = page pa. = paragraph → = should read

p.13, pa.2, line 16: *fuses* → *fuse*

p.13, pa.2, line 17: *undergoes* → *undergo*

p.20, pa.2, line 9: *rigourous* → *rigorous*

p.21, pa.2, line 14: *rigourous* → *rigorous*

p.23, line1: *only languages* → *types of languages*

p.36, pa.2, line1: insert *the* (before *most*)

p.47, line 1: *Even* → *Even though*. Replace fullstop with comma.

p.57, pa.2, line 1: *sociolinguisitc* → *sociolinguistic*

p.63, line 5: *is* → *are*

p.65, pa.2, line 16: *pie-piped* → *pied-piped*

p.67, pa.2, line 14: *so* → *less so*

p.85, pa.3, line10: add (after *pilot study*): *such as using unclear instructions, difficult words and expressions in test items, grouping similar sentences together, and running the test for too long*

p.90, pa.2, line14: after *investigated* add: *It should be noted that conversational data sampled for the informal of the styles, do not reflect truly informal speech style because the subjects will still be monitoring their speech to a fairly high degree. As compared with other types of data in the study, the conversational data are, nonetheless, much less formal in style.*

p.94, pa.2, lines 8-11: replace *The subjects were familiar with the task situation and would regard their participation as a practice opportunity for a real test; as a result, undesirable effects such as the Hawthorn effect could be reduced to minimum. with The subjects' familiarity with the task situation including methods used therein could help to reduce undesirable effects such as the Hawthorn effect to a minimum.*

p.107, pa.2, line 5: 349 → 350; 313 → 314

p.134, line 5: add a footnote after '*definiteness*' which reads: *This intuitive claim is drawn from the fact that that is used only in restrictive relative clauses whereas which is applicable in both restrictive and nonrestrictive relative clauses, and that that rather*

than which is preferred in some restrictive relative clauses with definite antecedents such as the first and the most important.

p.147, line 1: *filer* → *filler*

p.158ff., in the first line of each subsection (§4.3.1 through §4.3.6): *2.1.4* → *2.1.3.3*

p.172, lines 3-4: *25c: $x^2 = 13.04$, $df = 1$, $p < .001$* → *25c: $x^2 = 24.31$, $df = 1$, $p < .001$*

p.179, pa.2, line 12: *Whereas* → *However*,

p.179, footnote.54: add at the end: *Classified as passive voice are also those "covert passives" (reduced relatives with participles) as in "The dog ate the bone found by the boy".*

p.182, pa.2, line 4: *Whereas* → *But*

p.185, lines 11-15: delete the two sentences beginning with "*The overall...*" and "*That is...*" respectively, and replace with *Despite these marginal differences, these features do not seem to show amelioration over time, and they are therefore inadequate indicators of level of achievement.*

p.201, line 7: *discriminate* → *indiscriminate*

p.223, Appendix VI, last example: replace *The manager sacked the man who's negligence caused the accident.* with *The manager sacked the man which his negligence caused the accident.*

p.247: add the following bibliographic entries:

Song, J. J. (2001). *Linguistic typology: morphology and syntax*. Essex, England: Pearson Education Limited.

Tomlin, (1986). *Basic word order: functional principles*. London: Croom Helm.

**Syntactic Features of
the English Interlanguage of
Learners of English as a Second Language**

Thesis submitted in full fulfillment of the requirements for the degree of
Doctor of Philosophy

**Linguistics Program
School of Languages, Cultures and Linguistics
Faculty of Arts
Monash University
AUSTRALIA**

Mingjian Zhang M.A. (Melbourne)

January, 2003

Table of Contents

Abstract	vi
Statement	vii
Acknowledgements	viii
Abbreviations	ix
Chapter One Introduction	
1.1. Research questions	1
1.2. Underlying theoretical framework – typological approach and universals	1
1.3. Typological approach and second language acquisition	7
1.3.1. Typological approach in second language acquisition	7
1.3.2. Typological universals to be tested in the study	9
1.4. Significance of the research	9
1.5. Organisation of the thesis	10
Chapter Two Literature review	
2.1. Typology and universals	12
2.1.1. A brief history of typology	12
2.1.2. Approaches in typology	16
2.1.3. Crosslinguistic patterns regarding word order universals and grammatical hierarchies	22
2.1.3.1. Word order universals	22
2.1.3.2. Grammatical hierarchies	32
2.1.3.3. Some recently motivated implicational hierarchies/hypotheses	43
2.1.4. Summary	50
2.2. Second language acquisition	52
2.2.1. An overview: theory and practice in second language acquisition	52
2.2.1.1. Contrastive analysis	52

2.2.1.2. Error analysis and interlanguage theory	53
2.2.1.3. Monitor theory	55
2.2.1.4. Variation (and functionalism)	57
2.2.1.5. Universal approaches	61
2.2.2. Studies of interlanguage in question formation: a typological approach	68
2.2.3. Studies of interlanguage in relative clause formation	70
2.2.3.1. Relative clause formation in SLA: an overview	70
2.2.3.2. Relative clause formation in SLA: the Accessibility Hierarchy	72

Chapter Three Methodology

3.1. The study	83
3.1.1. Background of the study	83
3.1.2. Procedure of data collection and participants	84
3.1.3. Difficulties encountered in the sampling procedure	88
3.2. The tasks	89
3.2.1. Conversation – interview and role-play with cue card	90
3.2.2. Elicited repetition	96
3.2.3. Sentence combination	98
3.2.4. Grammaticality judgment	100
3.2.5. Free essay writing	104
3.3. Data processing	104

Chapter Four Results and analysis of findings

4.1. Results regarding implicational universals of <i>yes/no</i> and <i>wh</i> questions	106
4.1.1. Results of data	106
4.1.2. Interpretation of data and discussion	110
4.1.3. Summary	119
4.2. Results regarding the acquisition of relative clauses with reference to the Accessibility Hierarchy	119
4.2.1. Results from the repetition task	120

4.2.2. Results from the sentence combination task	124
4.2.3. Results from the grammaticality judgment task	133
4.2.4. Results from the conversation and subjects' written essays	137
4.2.5. Results of various RC-related non-standard forms	139
4.2.5.1. RC-related non-standard forms in different tasks	139
4.2.5.2. RC-related non-standard forms via English proficiency levels	150
4.2.5.3. RC-related non-standard forms via first languages	153
4.2.6. Summary	156
4.3. Results regarding other implicational universals	157
4.3.1. Results regarding the Clause Embedding Hierarchy	158
4.3.2. Results regarding the Subordinate Gap/No Gap Hierarchy	162
4.3.3. Results regarding the Hierarchy for Complementisers	163
4.3.4. Results regarding the Bridging Verb Hierarchy	166
4.3.5. Results regarding the Head Noun Phrase Hierarchy	168
4.3.6. Results regarding the Valency Completeness Hypothesis	170
4.3.7. Summary	174
4.4. Results regarding interlanguage features in naturalistic data	177
4.4.1. Overall results from the conversation and written essays	177
4.4.2. Results examined by different English proficiency level groups	183
4.4.3. Results examined by different first language groups	185
4.4.4. Summary	193

Chapter Five Summary of major theoretical findings

5.1. Findings regarding the acquisition of English interrogative questions	196
5.2. Findings regarding the acquisition of English restrictive relative clauses	197
5.3. Findings regarding other implicational universals	199
5.4. Summary	202

Chapter Six Concluding remarks

6.1. Methodological limitations of the study	203
6.2. Use of SLA data for the evaluation of typological universals	204

6.3. Some theoretical issues arising from the study that warrant future research	211
 Appendices	
Appendix I Table of Universals Tested, Data Collected and Methods Used in the Study	213
Appendix II Profiles of the Subjects	214
Appendix III Tasks for Data Collection	216
Appendix IV Counts of Instances in Features of Question Word Order	219
Appendix V Definitions of Features for the Repetition Task (A) and Individual Performances in the Repetition Task (B)	220
Appendix VI Definitions of Features for the Sentence Combination Task (A) and Individual Performances in the Sentence Combination Task (B)	223
Appendix VII Individual Performances in the Grammaticality Judgment Task (Items Related to Relative Clauses)	227
Appendix VIII Counts of Relatives Clauses in Subjects' Speech and Written Essays	229
Appendix IX Individual Performances in the Grammaticality Judgment Task (Items Related to Various Implicational Universals) (A & B)	230
Appendix X Aggregate Counts of Instances of Different Features in the Conversation Task	233
Appendix XI Aggregate Counts of Instances of Different Features in the Subjects' Written Essays	235
 Bibliography	 236

Abstract

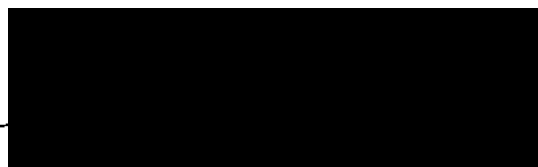
This study investigates syntactic features of the English interlanguage of adult ESL learners. The theoretical framework underlying the present study is mainly based on theories of linguistic typology and universals, particularly those of Hawkins (1999). Implicational universals such as the word order of interrogative question formation (Greenberg 1963), the Accessibility Hierarchy regarding relative clause formation (Keenan and Comrie 1977) and a number of processing-motivated implicational hierarchies/hypotheses in terms of filler-gap domains (Hawkins 1999) were examined in the study in order to test the extent to which the syntactic features of the English interlanguage of ESL learners can support the predictions made by these typological universals.

Data for the study was collected from about sixty international students studying at Monash University English Language Centre via a collection of written essays from the subjects as well as a number of tasks including a conversation (comprising an interview and a role-play), elicited repetition, sentence combination and grammaticality judgment. Data collection was carried out both cross-sectionally (from learners of different language-speaking backgrounds at different proficiency levels) and longitudinally (at two intervals over a year). Results of the study show that the syntactic features studied in the subjects' English interlanguage predominantly support the implicational universals under examination regardless of the English proficiency levels and first languages of the subjects, lending support to the implicational universals as valid predictors for the phenomena of second language acquisition. More significantly, the processing-motivated explanation for the implicational universals in terms of filler-gap domains (Hawkins 1999) provides a unifying account, which can address adequately both the regularities and irregularities of the interlanguage syntactic features under study. The theoretical implications of counterexamples found in this study are discussed.

STATEMENT

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

Signature



Date

3/2/2003

Acknowledgements

To begin with, I wish to express my heartiest gratitude to Dr Heather Bowe for her excellent supervision for this PhD thesis. From the very first day of my PhD candidature to the last date for the submission of the thesis, Dr Bowe has spent countless hours supervising me enthusiastically in every aspect of the research project ranging from the overall design and theoretical framework of the research to the style and wording of academic writing, never failing to encourage me to embark on the research project, the finding of which is reported in the thesis. During the three-and-half-year period of her supervision, Dr Bowe's regular supervision on my progress, enthusiastic encouragement, administrative support, critical comments and insightful suggestions all have lent greatly to the completion of the thesis. However, I am, alone, to blame for any weaknesses that reside in the thesis.

I also wish to express my thanks to Dr Keith Allan, who acted as my temporary supervisor during Dr Heather Bowe's sabbatical leave and provided valuable and insightful remarks upon my work. My sincere thanks to Olwen Beale, who contributed greatly to the data collection of the research by successfully administering the conversation task to all the subjects of the study. Thanks also to Ruth Rosen, Alice Sinclair, Noel Williams, Gary Peters, and Guangbo Chen, all of whom helped me in one way or other during the process of data collection, and also to the subjects of the study, who volunteered to participate in the research, and without whom the research would not have been possible.

Finally, special thanks to my wife, Xiaojun Che, and my son, Zhiyuan (Tom) Zhang; without their support and tolerance, it would have been very difficult for me to complete the thesis in due time.

Abbreviations

Det = determiner

DO = direct object

ESL = English as a second language

GEN = genitive/possessor noun phrase

IO = indirect object

L1 = first language

L2 = second language

N = noun

NP = noun phrase

OBL = oblique object

OCOMP = object of comparison

Prep = preposition

PP = prepositional phrase

RC (Rel) = relative clause

SLA = second language acquisition

SU = subject

V = verb

VP = verb phrases

Chapter One Introduction

1.1. Research Questions

The present research is a study of syntactic features in the interlanguage of ESL learners from the perspective of linguistic typology. The overall purpose of the research is to determine to what extent syntactic features of the interlanguage of ESL learners can be shown to reflect some of the language universals such as implicational universals and grammatical hierarchies found in the linguistic typology. Put more specifically, the research is to investigate which syntactic features in the interlanguage development of ESL learners in a second language setting are in accordance with the patterns of typological universals and which are not and why. The research aims to provide a framework for describing and explaining the features of interlanguage development of ESL learners; however, implications for linguistic typology will also be addressed. The research questions for the study are:

- 1) How are adult ESL learners developing their interlanguage at the syntactic level?
- 2) Are there any syntactic features in the interlanguage of ESL learners that reflect the typological universals reported in the literature of typology?
- 3) Are there any syntactic features in the interlanguage of ESL learners that violate the constraints of typological universals?

1.2. Underlying Theoretical Framework – Typological Approach and Universals

Modern linguistics sees a number of different schools of thought addressing the fundamental question, "What is a possible human language?"; among them are two major approaches - Generative Grammar and Language Typology, both of which are engaged in uncovering the universality of language. Though the two approaches share some commonalities in trying to delimit the universal constraints on language structure (particularly at the syntactic level) and to reach the abstraction of universals, they are

diametrically opposed to each other in their basic underlying assumptions and methodologies.

The generative approach represented by Noam Chomsky and his fellow researchers claims that all human beings are genetically endowed with an innate language faculty. This faculty contains a 'Universal Grammar' (UG) that enables the child to learn rapidly any complex and mature grammatical system in the world and the core of this innate UG must be embedded in any human language. Under this rationalist view of language, the universality of language is deductively sought within only a limited number of languages such as English. The UG thus formulated is consistently explained in terms of the abstract formal constructs of syntactic structures in line with its formalist autonomy thesis (independent of semantic and functional considerations) no matter what evolutionary phase it is in (e.g. Chomsky 1965, 1981a and 1993 representing the classical generative approach, principles-and-parameters approach and minimalist approach respectively).

In contrast, the typological approach represented by Joseph Greenberg and subsequent researchers holds that the universality of languages can only be discovered through the comparative examination of a large variety of world languages. Thus formulated crosslinguistic generalisations represent universal constraints on human language which could otherwise be missed through examining a single language or a few languages. Accordingly, empirical methods are applied in sampling a variety of languages, universal patterns and constraints discovered in the crosslinguistic data focus on surface form and meaning, and explanations sought for the universals are multidimensional including language-external factors such as functional factors as well as language-internal factors such as structural and diachronic factors.

There has been some increased interest in integrating typology into generative grammatical theory (Greenberg 1991b; Fukui 1995), despite the fact that there have been some attacks on the typological approach to universals by generativists (e.g. Coopmans 1983, 1984). Hawkins (1988b) and Greenberg (1991b) argue that the two approaches are complementary, each having its own strengths. Shibatani and Bynon (1995) further

suggest that these two approaches are converging in the sense that the typological assumptions have influenced the significant shift in the orientation of UG to the 'principles-and-parameters' approach, which in turn has had a greater impact on typological studies dealing with entirely unrelated languages. In fact, the typological approach is argued by Hawkins (1988b) to have certain advantages over the generative approach in documenting large scale variation across numerous languages, revealing crosslinguistic patterns therein and identifying interacting explanatory principles behind those patterns. In view of the nature of the present study, in which syntactic features of the interlanguage of ESL learners from different language speaking backgrounds are under examination, the typological approach to universals is adopted as the basic theoretical framework that underlies the present study.

Modern syntactic typology is taken to represent the tradition of cross-linguistic study of morphosyntactic¹ properties beginning with Greenberg's word order universals and continued largely by American linguists (Croft 1995), hence often referred to as the Greenbergian linguistic typology (Croft 1995; Song 2001). Since Greenberg's original work on word order, typologists have developed four basic types of typological universals – implicational universals, markedness, grammatical hierarchies, and prototypes.

An *implicational universal* is a fundamental type of universal characteristic of most typological research. It captures a pattern of co-occurrence regularities between two language parameters in the form "If a language has P, then it also has Q". Another important type of typological generalisation is *markedness*, which "is a property of a grammatical category such that it displays one or more of a cluster of grammatical asymmetries cross-linguistically". (Croft 1995: 106) It is different from the Prague school notion of markedness in that the former is a crosslinguistic generalisation applicable to function as well as form and with an emphasis on behavioural and frequency criteria. A *grammatical hierarchy* characterises a pattern of crosslinguistic variation in the form of

¹ 'Morphosyntactic' instead of 'syntactic' is used here because of the blurred division between morphology and syntax in modern syntactic typological studies (Croft 1995: 85).

ranking of members within the same grammatical category. It is more complex than the former two in that it can be derived from a chain of implicational universals or it can be seen as a series of markedness patterns in which relative rather than absolute values of markedness are adopted. A typological *prototype* is "an ideal example of a category" (Whaley 1997: 289), which characterises a pattern of crosslinguistic variation in the form of a cluster of grammatical values defined in other categories. An 'ideal' grammatical form should possess all these grammatical values, hence prototypical, while a grammatical form lacking one or more of these grammatical values is prototypically marked or even loses its category membership.

As can be seen, these types of universals do not stand alone all by themselves; rather they are interrelated to one another. For instance, markedness underlies the various grammatical hierarchies and prototypes and their interactions, and the marked-unmarked relationship of singular and plural can be captured in the implicational universal: "If the plural is expressed by the absence of a morpheme, then so is the singular." (Croft 1990: chapters 4-6)

The major crosslinguistic patterns and their interactions have been applied widely by typologists to address a range of morphosyntactic phenomena across languages. In word order typology, emphasis has moved from simple implicational universals which prove to be very effective in capturing patterns of word order in Greenberg (1963), to more complex and exceptionless implicational universals such as Hawkins (1983) and statistically significant implicational universals such as Dryer (1992). The issue of morpheme order is also dealt with in light of implicational universals (Hawkins and Gilligan 1988). Markedness patterns can be found in various grammatical categories, including hierarchies (Greenberg 1966) and prototypes (Croft 1990). For instance, in discussing the distinction between direct-indirect object system and primary-secondary object system, Dryer (1986) shows that direct and primary objects are less marked than indirect and secondary objects by the frequency as well as structural and behavioral criteria of markedness. In comparing the relative clause forming strategies used by a number of languages, Keenan and Comrie (1977) formulate a grammatical relations

hierarchy of relativisation, i.e. 'accessibility hierarchy' (AH: subject < direct object < indirect object < oblique); in the hierarchy, all languages must be able to relativise subjects, all the possible relativised NP positions in a language should be contiguous, but languages vary at different cut-off positions down the hierarchy. Hopper and Thompson's (1980) study of transitivity is a well-known application of prototype analysis to grammatical categories across grammars. They propose a set of prototypical transitive features across related grammatical categories; typologically, these features are all concerned with the transitivity of a clause, but a particular language has conventionalised a particular set of features that affect the transitivity of a clause.

Similarly, in the present study, some other typological universals as well as the implicational universal regarding question acquisition order and the accessibility hierarchy regarding relative clause formation are examined and the interlanguage data are addressed in terms of these typological universals and their interactions.

The crosslinguistic patterns as well as the basic typological universals discovered from the studies of the world's languages have formed the typological proper of defining and limiting possible variation in human languages. However, these typological universals are in large part observational and descriptive, and are therefore low-level generalisations; but they provide the data that a theory of language must account for. With the discovery of universal patterns across languages, typologists have been seeking higher-level generalisations, i.e. deeper principles that underlie the universals, from either internal language system or considerations outside of the language system. Since the 1970s, modern syntactic typology has evolved into a functional-typological approach as a theory of grammar, which seeks deeper principles from external motivations to account for crosslinguistic patterns (Croft 1990, 1995).

Externally motivated explanations for observed universals have been sought in related areas of functions such as economy, iconicity, discourse, perception-cognition and processing. The economic motivation or 'economy' is an important principle that the more frequently used expressions tend to have fewer morphemes. This motivation

pervades grammatical expression and offers a plausible explanation for most markedness patterns and some other aspects of grammar and typology (Haiman 1985). The iconic motivation or 'iconicity' is another important principle that the form of linguistic expressions reflects in some way the real-world structure of experience. This semantically-oriented account of grammatical structure is sought in isomorphism (one form, one meaning) (Haiman 1980, 1985), and the relations between conceptual distance and linguistic distance (Givón 1980; Haiman 1985; Bybee 1985). A discourse-based external explanation is the communicative motivation, the principle that language is capable of expressing all the conceptual structures via some grammatical means. This motivation underlies the typological conspiracies, in which logically independent grammatical processes 'conspire' differently in different languages to achieve the same effect (Croft 1995). A conspiracy is observed in some languages (e.g. Bantu languages: only subjects and objects can be relativised) by Givón (1979) between accessibility of NP relativisation and 'promotion' of noun phrases lower in the AH such as oblique object to a higher position such as object: in these languages, virtually any noun phrase can be relativised in the light of conspiracy. The perceptual-cognitive motivation is particularly exemplified in the area of lexical semantics (Whaley 1997). The implicational hierarchy of basic colour terms in language (white-black > green-yellow > blue > brown) proposed by Berlin and Kay (1969) shows that there exist universal constraints on the types of possible basic color lexicons. These constraints are perception-grounded arising from the structure and function of the visual system (Kay and McDaniel 1978).

Finally, the processing motivation is one of the most important principles that have been used to account for crosslinguistic patterns. The fundamental assumption of this principle is that all humans share similar processing constraints which lead them to avoid structures that are hard to understand and produce and to favour structures that facilitate rapid comprehension and production, and that consequently, these processing constraints are imposed on the structure of their grammars. The processing-motivated explanation underlies Dryer's Branching Direction Theory (1992) and particularly a series of works of Hawkins. Hawkins (1985, 1988) appeals to processing evidence and principles in psycholinguistics in addressing crosslinguistic regularities regarding suffixation

preference and co-occurrence of affixation and basic word orders. Later, Hawkins (1990, 1993, 1994) develops a global principle – the Early Immediate Constituents (EIC) as a major processing principle in his parsing theory of word order universals. Recently, Hawkins (1999) has furthered his processing explanation for crosslinguistic patterns in terms of filler-gap dependencies. The EIC and Hawkins' (1999) recent processing account can be regarded as both internally- and externally-motivated explanations, i.e. explanations that incorporate insights from generative grammar as well as from psycholinguistics.

To conclude, the functionally-oriented typological approach to universals is “an ‘approach’ to the study of language that contrasts with prior approaches, such as American structuralism and generative grammar. In this definition, typology is an approach to linguistic theorizing, or more precisely a methodology of linguistic analysis (Croft 1995: 86). Based on this typological approach in general and Hawkins (1999) in particular, the present study is a test of the applicability of some implicational universals and hierarchies in terms of filler-gap dependencies in second language acquisition. Validating the adopted theoretical framework for the research is a detailed literature review in chapter two.

1.3. Typological Approach and Second Language Acquisition

1.3.1. Typological Approach in Second Language Acquisition

It is argued by some typologists that the typological approach, which is viable in investigating linguistic universals across languages, should also hold for language acquisition. Hawkins holds that implicational universals for diachronic as well as synchronic predictions should be manifested in language acquisition prediction as well; “in language acquisition as in historical change, the consequent may be acquired first, or both may be acquired simultaneously, and all that we can rule out is the acquisition of the antecedent prior to the consequent, since there are no language of this type currently attested.” (1990: 99) Comrie has made the statement that, apart from the investigation of a wide range of primary languages, “Another area where one can study spontaneous innovation is in child language... Similarly, one could study the acquisition of a second

language, to see if any universals are mirrored in its acquisition process, especially in cases where those universals are not the subject of direct evidence in either the native language or the target language." (1981: 222)

Eckman describes the viability of the approach in SLA on two grounds: "First, its claims about SLA are readily testable. And second, it defines a fruitful program of research." (1993: 64) By the 'testability' of the approach, the universal generalizations formulated as implicational statements can be tested empirically on secondary languages including interlanguages, and can be accounted for by "Whatever linguistic theory ultimately proves defensible" (Eckman 1993: 65). And by the 'fruitfulness' of the approach, "The intent of typological generalizations is to characterize the range of possible variation in human languages" and "an obvious relationship can be stated between typological generalizations and the explanation of facts about secondary languages" (Eckman 1993: 66). Concerning secondary-language acquisition, typological universals are claimed to be employed to make predictions about degree of difficulty (Eckman 1977), language transfer (Gass 1979), and order of acquisition (Hytenstam 1984) with respect to various target structures. From within an essentially Chomskyan approach, Yip and Matthews (1995) also make an attempt to include typological issues such as topic-prominence into theories of interlanguage.

Eckman (1991) postulated the interlanguage structural conformity hypothesis (ISCH) stating that the universal generalizations that hold for the primary languages also hold for interlanguages. The ISCH was tested against two principles of interlanguage phonetic structure, the fricative-stop principle and the resolvability principle, and was strongly supported by the results of his study. And again this hypothesis was confirmed by Carlisle (1997) in a recent study that examined how frequently native Spanish speakers modified English two- & three-member onsets.

In a different study, Hawkins (1987) makes implicational universals as explicit as predictors of language acquisition, through his insightful illustration of implicational universals based on critiquing Jakobson's phonological universals and against some

acquisition data. The universals can be used to predict order of acquisition in terms of production, comprehension and direction of substitutions quantitatively. Even more significant is Hawkins' claim that not only are these universals practically falsifiable, they may provide explanation for the route and nature of language acquisition as well.

The fact that the typological approach has been effectively applied in addressing issues in second language acquisition shows that the field of second language acquisition is a valid field for testing language universals. Testing these universals against interlanguage may also help to further understand the characteristics of second language acquisition, language universals and human language capacity.

1.3.2. Typological Universals to Be Tested in the Study

The present study tests the extent to which the implicational universals can be applied to explain the interlanguage of ESL learners in the following aspects:

- a. *yes/no* and *wh*-question formation;
- b. relative clause formation; and
- c. filler-gap dependencies in various structures.

Refer to **Appendix I** for details.

1.4. Significance of the Research

The typological approach is adopted in this piece of research and the typological universals mentioned above are tested against interlanguage data of ESL learners. The adoption of the typological approach for this study is justified on the following grounds.

First, it is believed that universal properties discovered from the study of primary languages should also hold for secondary languages (e.g. Comrie 1981; Hawkins 1990; Greenberg 1991; Eckman 1993). Since typological universals reflect linguistic patterns and variations across primary languages, the logic of testing these universals in the secondary forms of language such as interlanguages is clear: if they do hold for interlanguage data, then the universality of these universals can be regarded as more deeply rooted, thereby confirming the identical functioning of human language faculty in

both first and second language acquisition. In effect, the modern field of typology, as Greenberg puts it, "has been able to produce concrete results in the form of generalizations, typically implicational, which provide hypotheses that up to now have played a major role in second language acquisition studies." (1991: 37), and there have been quite a few studies in which universals have been used to address phenomena of second language acquisition (see 3.1).

Second, this approach implies a dynamic view of universals in relation to the process of language acquisition as opposed to language transfer; "It all depends on our conception of universals, which I take not to be a fixed and static set of principles, but rather relational and diachronic. The first is involved in the notion of markedness, always a relation and a hierarchic one among linguistic properties. The second is inherent in the very notion of interlanguage as a process which takes place over time." (Greenberg 1991: 41) In fact, all the implicational universals involve, to varying degree, relational hierarchies, which predict not only possible versus impossible co-occurrences of language but the direction of language development as well. The dynamicity of the approach can therefore best capture the dynamic nature of interlanguage development.

All in all, the proposed research using a typological approach attempts to reveal how adult ESL learners develop their interlanguage at syntactic level in an ESL setting, which characteristics of their interlanguage are universal, and which are not and why, thereby providing a framework for describing and explaining interlanguage development for ESL. It was expected that the research would make significant contributions in two respects: to reveal the extent to which a typological approach can be applied to the area of second language acquisition and the extent to which typological universals can be manifested in interlanguage data, and to deepen the understanding of the process of second language acquisition.

1.5. Organisation of the Thesis

This thesis consists of six chapters. The present chapter has introduced the basic research questions, the underlying typological framework for the research, the viability of

typological approach in second language acquisition, and the significance of the research. Chapter Two presents a selective literature review relating language typology, second language acquisition, and implicational universals and hierarchies relevant to this study. Chapter Three describes methodological issues of the research including the design of the study, data gathering, informants, data processing and difficulties encountered. Chapter Four presents a discussion of the results of the study regarding question formation, relative clause, and some other implicational universals. Chapter Five summarises major theoretical findings of the study. Chapter Six discusses the methodological limitations of the study, the use of L2 data for the evaluation of typological universals, and related areas for future research.

Chapter Two Literature Review

This chapter reviews relevant literature in both typology and second language acquisition. Section One (2.1) includes a brief history of typology, different typological approaches, and crosslinguistic patterns regarding word order universals and grammatical hierarchies. Section Two (2.2) consists of an overview of theories of second language acquisition and interlanguage and an examination of the SLA studies concerning *yes/no* and *wh* questions and relative clause formation with particular reference to the Accessibility Hierarchy.

2.1. Typology and Universals

This section begins with a brief account of the evolutionary process of typology from morphological studies of languages in the nineteenth century to modern linguistic typology with the emergence of Greenbergian word order universals (2.1.1), illustrates different theoretical orientations and practical methodologies associated with a variety of approaches employed in modern linguistic typology (2.1.2), then reviews modern typological studies concerning word order universals, grammatical hierarchies and some recently motivated implicational hierarchies/hypotheses (2.1.3), and ends with some summarising remarks (2.1.4).

2.1.1. A Brief History of Typology

The morphological typology of the nineteenth and early twentieth centuries focused on a classification of language types across languages on the basis of their morphological characteristics. The original classification was formulated by Friedrich von Schlegel, who divided world's languages into two major types: affixal and inflectional. The former used the simple combination of morphemes as a means of forming grammatical relations, while the latter demonstrated a more complex type with phonological alterations of morphemes in combination. Later, his brother, August Schlegel found that the two types could not include languages like Chinese, and added a third type-languages with 'no structure', the type that has no affixation or inflection but relies on word order to express grammatical relations. Another well-known German linguist, Wilhelm von Humboldt again found that even these three types could not generalise all the languages in the

world. In analysing sentences of world's languages, he noticed that some languages such as the Eskimo and some other American Indian languages treated the verb and the object as a whole word. As a result, he added a fourth type – 'incorporating' or polysynthetic language where the principal or the whole structure of the sentence was embedded in a single word in which grammatical relations were shown through combination of bound morphemes.

The classical formulation of morphological classification of language types came from a versatile historical linguist, August Schleicher. Influenced by Haiguer's philosophical idea, August Schleicher believed that language consists of form and meaning and there does not exist a language form without meaning. Accordingly, languages were classified into three types: isolating, agglutinative and inflectional and any language was taken to belong to one of the three types. Under this classical formulation of types, isolating languages refer to those where grammatical forms do not affect grammatical meaning at all like Chinese, in other words, grammatical relations are mainly realised through different word orders of unchanged root words as there are no affixes and inflections in these languages. Agglutinative languages are those where language units include both grammatical forms and meaning like Turkish, that is, these languages use free combination of affixes onto the unchanged roots to express grammatical relations with each affix denoting a single grammatical category. Inflectional languages denote those where grammatical forms and meaning are synthesised through inflexion such as Latin and Greek, namely, in these languages, grammatical relations are expressed through affixes which often fuses together more than one grammatical categories into a single morpheme and often undergoes major phonological alterations when combined with roots or stems.

As can be seen, the typological studies of language at that time focused on the structural properties of morphology only, which reflects the prevalent philosophical view of language of the time, namely, any human language is thought to have its own abstract organic unity. Under this view, the formal aspects of any language were thought to reflect its inner character and morphology was regarded as the essential aspect of language;

therefore, the study of morphological types was thought to best reveal the whole inner organic nature of language. As a result, the early typology flourished in classifying language types on a single parameter of morphology.

At the beginning of the twentieth century, Ferdinand de Saussure's "Course in General Linguistics" ushered in a new era in the field of language studies, often referred to as the beginning of modern linguistics. Some fundamental concepts in Saussure's theory such as *langue* versus *parole*, synchronic versus diachronic study of language, syntagmatic and associative relations, have laid the theoretical foundation for general linguistics and lent differently to almost all schools of thought particularly structuralism in the twentieth century. Under the structuralist movement, although the view of language as an organic whole still persisted, the idea of synchronic study of discrete parts of language along with rigidity in methodology altered the view of morphological typology of languages. Instead of focusing solely on the whole morphological types on languages, it was now possible and necessary to study parts of language and to make a typological classification of various features and different parts across languages in different ways. Edward Sapir (1921) revised morphological typology along these lines.

Sapir's contribution to typology was significant in that the typological classification of languages was not to be confined to the whole structural types of languages and to morphology solely but to be extended to other types and areas of languages as well. For example, using one parameter - the number of morphemes per word, Sapir (1921) classified languages into three types: analytic, synthetic and polysynthetic; using another parameter - the degree of phonological alteration of morphemes in combination, he classified languages into four types: isolating, agglutinative, fusional, and symbolic. Incorporating a semantic component, he also postulated a complex typology in terms of the ways that languages might vary in expressing four types of concepts: a) concrete (basic lexicon), b) derivational (forming new words via affixes), c) concrete relational (indicating grammatical relations but with some concrete meaning) and d) pure relational (denoting grammatical relations only). Languages can be classified into four different types according to the combination of types of concepts they adopt: A. simple pure

relational (a, d), B. complex pure relational (a, b, d), C. simple mixed relational (a, c, d) and D. complex mixed relational (a, b, c, d).

Based on the examination of the 19th century morphological typological studies including Sapir's work, Greenberg (1954) used quantitative indices to typologise quite a number of parameters for morphemes, words and grammatical relations, which furthered morphological typology with a solution to the problem of arbitrarily forcing a specific language into a morphological type by traditional morphological typology and with a more sophisticated and better articulated model of morphological types. His work along the line of morphological typology, together with the work of previous typologists, paved the way for the emergence of what is often referred to as modern linguistic typology.

On the basis of holistic morphological typology in the nineteenth and early twentieth centuries, modern linguistic typology has shifted its focus to "the practice of partial typology, where specific constructions and grammatical phenomena, such as word order, case-marking patterns, relative clauses, passives, causatives, are examined, typologized, and classified" (Shibatani and Bynon 1995: 9).

Modern linguistic typology or modern syntactic typology in particular, is closely associated with the work of Joseph Greenberg, who organized the significant Dobbs Ferry Conference on Language Universals in New York in 1961 and edited the conference papers into a ground-breaking book entitled *Universals of Language* (Greenberg 1963). Greenberg's formulation of implicational universals of morphology and word order in his own paper in the volume ("Some universals of grammar with particular reference to the order of meaningful elements"), has founded the source of modern study of typological universals. Over the last forty years, the field has developed rapidly to include many research institutes (e.g. Max Plank Institute) and research centres (e.g. Cologne), different approaches (to be dealt with in next section), and a huge amount of published literature in both individual language studies (e.g. Bowe 1990) and studies across languages (e.g. Comrie 1981; Hawkins 1983; Dryer 1988).

2.1.2. Approaches in Typology

Apart from the modern syntactic approach from the Greenbergian tradition discussed in section two of chapter one, research in language typology now embraces a range of other approaches with different theoretical orientations and practical methodologies including the Leningrad Typology Group, the Cologne UNITYP Group, the Prague Typology Group, the Paris RIVALLC Group and diachronic approach (Shibatani and Bynon 1995).²

The Leningrad (now St Petersburg) Typology Group was founded in the USSR (now Russia) in the early 1960s with Aleksandr A. Xolodovic (1906-77) the inaugural head of the group. Over several decades, the Leningrad Typology Group has provided detailed descriptions of well-defined, specific grammatical constructions both in single languages and across languages with their 'collective typology' method (Nedjalkov and Litvinov 1995). Researchers under this team-work approach assume that typological descriptions rather than universals are of primary concern, and that theoretical underpinnings subordinate to such descriptions are only behind their research questionnaire because they believe that "the function of theory is to ask questions which empirical investigation ought to answer." (Nedjalkov and Litvinov 1995: 257) In consequence, this description-oriented typology group has produced a good amount of collective publication covering a variety of grammatical constructions such as causative, passive, resultative and iterative in great depth.

The Leningrad Typology Group typology differs from the Greenbergian tradition of linguistic typology in the following respects. First, the former focuses on typological description as "the ultimate goal of this research" (Nedjalkov and Litvinov 1995: 259), while the latter aims to delimit the possible range of language variation in terms of universals or universal constraints. Furthermore, in the Greenbergian linguistic typology, language universals research is based on typological description and "is merely the next logical step forward" (Song 2001: 342) from such description. Second, this group's lack

² There may be other approaches to linguistic typology such as the research group in Sweden, the EUROTYP (Typology of Languages of Europe) project in Europe and typological studies at the Research Centre for Linguistic Typology in Australia (Song 2001). limitations of space and focus on (Shibatani and Bynon 1995) preclude discussion of them here.

of concern for explanation may be due to the exclusion of universals in typological research, inherent restraints of the 'collective method in typology' and their reliance on form (structural features) rather than on function (Song 2001), whereas the Greenbergian linguistic typology seeks explanation of typological universals from functional as well as language internal motivations. Finally, the Greenbergian linguistic typology is concerned with methodological issues such as statistical distribution of types as a significant index in explanation and sampling methods (Whaley 1997) to reduce areal and genetic biases, which are generally out of the Leningrad Typology Group's consideration.

In 1972, Hansjakob Seiler founded a research project on language universals and typology known as the Cologne UNITYP Group, which is still operative and very productive today. Influenced by the Humboldtian view of language as an activity rather than a product, the UNITYP typologists treat language as a problem-solving system, and "The main endeavour of UNITYP thus far has been firmly to substantiate the view that the essence of language is the process, and not the thing" (Seiler 1995: 299). To identify and reconstruct such a process in typological studies has thus become the main research objective; a complex framework of cognitive-conceptual structure and linguistic encoding is postulated to obtain the objective. In the framework, the dimensions of investigation are deductively posited and construed as cognitive-conceptual domains including nomination, concomitance, determination, possession, apprehension, participation, situation and localization, and each dimension consists of a number of sub-dimensions identified inductively through analysing relevant data from a single language or from a variety of languages. Thus that cognitive-conceptual dimension is represented by a continuum of these sub-dimensions, out of which each language has a particular choice via the 'mental operation'. The mental operation is defined as comprising three functional/operational principles – 'indicativity' (inherent relationship), 'predicativity' (established relationship), and 'iconicity' (relationship of similarity) (Seiler 1995: 278). Only the three principles resembling the 'process' are regarded as universal, the complex operations of which link the cognitive-conceptual domains (invariant unity) to the specific forms and meanings of the world languages (variant diversity).

Both the Greenbergian linguistic typology and the UNITYP typology are functionally oriented. However, the former appeals to function in their externally motivated explanations for typological universals (e.g. economy and processing), while the latter's functional stance is couched in their view of language as a 'problem-solving system' whereby domains of investigation are conceptualised in cognitive-conceptual terms (e.g. possession and determination) which in turn are expressed by linguistic means. Moreover, the universality sought by the UNITYP typology is also closely related to the mental operations of the three principles (indicativity, predicativity, iconicity) not to the empirical crosslinguistic generalisations, in which the universals research of Greenbergian typology mainly lies. Finally, similar to the Leningrad Typology Group, the UNITYP seems to be less concerned with the issues of distribution of types and language sampling (Song 2001), which the Greenbergian typology has always taken seriously and developed rigorous methods for.

The Prague School is one of the major schools of thought in modern linguistics. The theory of markedness developed by Praguian linguists Nicholas Trubetzkoy and Roman Jakobson has influenced both generative grammar and modern linguistic typology, which have adapted markedness in their respective theories in different ways (Croft 1990). The Prague School Typology is mainly associated with Czech linguists Vladimír Skalicka and Petr Sgall, who are, in sharp contrast to typologists in other approaches such as the Greenbergian tradition, the Leningrad Typology Group and the UNITYP, ardent proponents of holistic typology (Sgall 1995; Song 2001). Based on the traditional holistic typology, Skalicka classified his holistic typology into five types – agglutinative, inflectional, isolating, polysynthetic and introflexive with each type possessing a number of basic properties from different grammatical domains. Central to this holistic typology are the relations of 'favourability' of one property to another between properties. While Skalicka's favourability is symmetrically bilateral in that if property X is favourable to property Y, then vice versa indicating the co-occurrence of the properties, Sgall argues for an asymmetrically unilateral favourability which means that if property X is favourable to property Y, it is not always the case vice versa (Sgall 1995). By this unidimensional relation of favourability, "it is possible to substantiate the claim that each

of the cited basic properties is favourable to the other features of the given type, concerning not only morphemics, but also other layers of language structure" (Sgall 1995: 65). The holistic typology in which formal properties of grammatical values interrelated in terms of unilateral favourability cover not only morphology but phonology and syntax as well, is claimed to be better than other partial typologies in terms of degree of holism (Sgall 1995: 71-5).

Much improvement as there is in the holistic Prague School Typology as compared with the traditional holistic morphological typology, the mainstream of contemporary typological studies is still dominated by partial typology as practiced by many other approaches. Like the Leningrad Typology Group, the Prague School Typology is not concerned with the search for universals, and therefore is lacking in explanation. Moreover, many issues concerning Sgall's holistic typology such as the unclear predicative power of the basic property in the relation of favourability, the arbitrariness of drawing distinctions between the five types, the difficulty of classifying languages into one of the five types even in terms of prototypical and marginal, lack of systematic data in support and language sampling methodology, are all called into question (Song 2001: 253-5).

In France, the Paris RIVALLC Group (*Recherche interlinguistique sur les variations d'actance et leur corrélats*) was found in 1984, dealing exclusively with the issue of transitivity in terms of actancy variation (grammatical relationships between the verbal predicate and the main noun phrases in verbal sentences). They believe that universality of language lies not in its grammatical categorisation but in its "conditions governing the functioning of language" (Lazard 1995: 179). While actancy and actancy variation are defined in morphosyntactic formal terms such as verbal marking for cross-referencing, specific morphological markers, sentence position and specific syntactic transformations, those universal prelinguistic conditions, i.e. invariants, are semantically and pragmatically motivated consisting of relationships showing possible combinations of an actancy variation with its formal correlates. Thus-pursued typological study can "contribute to a better understanding of the relationships between, on the one hand,

processes in the real world as reflected by the human mind (semantics) and constraints of communication (pragmatics) and, on the other hand, the functioning of language with its own internal dialectics, its relative inertia, and the unequal plasticity of its different components" (Lazard 1995: 204).

As compared with the Greenbergian linguistic typology, the RIVALC typology is more 'partial' and formal in that it deals exclusively with the issue of actancy in greater depth and in pure formal morphosyntactic terms, even though its underlying assumption is meaning/function-oriented. While the Greenbergian universals arise from empirical generalisations, the RIVALC's invariants try to capture relationships based on analysis and comparison of actancy variations within and between languages. However, like the above-mentioned approaches, the RIVALC typology too shows lack of concern for seeking deeper explanation for the typological invariants and their methodology similarly focuses on the rigorous way of forming an actancy framework rather than on the distribution of types and language sampling.

Since the 1970s, diachronic approach has again played an important role in typological study, yet different from the traditional morphological typology in its reference to a partial typology (e.g. word order) and its role in linguistic explanation. Greenberg has long been concerned with the diachronic as well as synchronic typological research even when the structuralist synchronic studies in the academic circles still prevailed. Having first tried to distinguish the focuses of the genetic comparison method (for genetic classifications) and the typological comparison method (for typological universals), Greenberg has merged them into a more comprehensive comparison method (for language changes), 'the processual approach' in his data analysis "where it was seen that the true regularity lies in the dynamic tendencies, that is, the diachronic universals" (Greenberg 1969: 113). In this approach, the states of synchronic typological generalisations represent stages of a diachronic process, via which those exceptions to synchronic universals (such as a combination of SOV and N-GEN word orders) can be explained as unstable transitory stages (Greenberg 1969). Therefore, in generalization of internal language structure, Greenberg (1979) proposes levels of generalisation and

shows that diachronic typology represents the highest level over the levels of synchronic description and explanation, and that generalizations over language processes must ultimately replace generalizations of language states. In his "state-process" model of diachronic typology, whereby languages may be classified in probabilistic terms into various typological states of a process of dynamic development with transitions in between, Greenberg argues that it is through exploring this diachronic process that one can "provide an explanatory key for an enormous range of superficially disparate linguistic phenomena" (1995: 164).

To sum up, different approaches in modern linguistic typology, albeit varying in theoretical orientations and practical methodologies, have something in common that differentiates themselves from the typology in the nineteenth and early twentieth centuries. First, modern typologists classify languages using several parameters on which languages vary, rather than a single parameter-the morphological structure of words, and they also try to show how these parameters are related to one another. Even the holistic typology of the Prague School covers phonology and syntax as well as morphology. Second, modern typologists have developed a variety of more sophisticated means (as manifested in the illustration of different approaches above) in their typological classifications and/or universal generalisations. They also try to seek deeper explanation (though varying in degree of interest from each approach) for these classifications and/or generalisations from both intralinguistic and external factors and from both synchronic and diachronic perspectives (particularly from the Greenbergian linguistic typology). And third, more rigorous methodologies such as type quantification and language sampling methods have been developed in typological and universal studies (again particularly from the Greenbergian linguistic typology). In a word, all these factors not only distinguish modern linguistic typology from the traditional morphological typology, they contribute, in a true sense, to establishing modern linguistic typology as "an approach to linguistic theorizing, or more precisely a methodology of linguistic analysis" (Croft 1995: 86).

2.1.3. Crosslinguistic Patterns regarding Word Order Universals and Grammatical Hierarchies

As has been illustrated in section two of chapter one and the above section, the Greenbergian linguistic typology is characterised as having developed different types of typological universals (implicational, markedness, hierarchies, prototypes and diachronic) and externally- as well as internally-motivated explanations for them. This section presents a selective review of the crosslinguistic patterns regarding word order universals (2.1.3.1), grammatical hierarchies (2.1.3.2), and some recently motivated implicational hierarchies/ hypotheses (2.1.3.3), which are relevant to the present study.

2.1.3.1. Word Order Universals

Modern linguistic typology or Greenbergian linguistic typology began with Greenberg's seminal paper on word order universals (Greenberg 1963). The emergence of the word order typology is significant in that not only does it mark the inception of modern linguistic typology with its shifted focus from 'holistic' to 'partial' typological studies, but it proves to be a productive ground-breaking area which has inspired many typologists and has generated empirical and theoretical insights into the nature of human language.

On the basis of a 30-language sample, Greenberg (1963) examined word order patterns systematically and formulated 45 separate typological universals. There are two types of universals – unrestricted and implicational; the former being a generalisation of the property shared by all languages on one parameter (e.g. the order of subject and object) with the latter on two or more parameters (e.g. subject-verb inversion and position of question word/phrase). Greenberg's first universal is a case of an unrestricted one, which states (1963: 61):

- (1) In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object.

This unrestricted universal asserts that only languages with subject before object are attested in world's languages, while the other logically possible types – languages with object before subject are not attested (or extremely rare). Accordingly, the basic word orders dealt with in Greenberg's universals are VSO, SVO, and SOV respectively (1963).

More significant, however, are Greenberg's implicational universals, most of which (twenty-five counted by Hawkins 1983: 22) reveal correlated dependency between two apparently logically independent parameters – basic word order (VSO, SVO or SOV) and some other grammatical constituents (e.g. adposition). Universal 3 (Greenberg 1963: 62), for example, is such a case in point in which the basic word order VSO correlates with the distribution of prepositions:

- (2) Languages with dominant VSO order are always prepositional.

Another implicational universal, which pertains to word order in questions and is to be examined in this study, is Greenberg's Universal 11 (Greenberg 1963: 65) as presented in (3):

- (3) Inversion of statement order so that verb precedes subject only occurs in languages where the question word or phrase is normally initial. This same inversion occurs in yes/no questions only if it also occurs in interrogative word questions.

This universal in fact contains two implicational universals (Eckman et al 1989) or an implicational hierarchy, which can be expressed as 4a and 4b respectively in (4):³

- (4) a. *Yes/No Inversion* > *Wh Inversion* & *Wh Inversion* > *Wh Fronting*
 b. *Yes/No Inversion* > *Wh Inversion* > *Wh Fronting* ('>' means 'implies')

³ I realise that the term *fronting* is a relic of transformational grammar, and most modern theories do not assume movement has actually taken place (i.e. initial *wh* word/phrases analysed as sentence-initial base-generation as opposed to movement). However I find it a convenient term to use here, partly because it has been used by other researchers such as Eckman et al (1989), and partly because alternative expressions are rather clumsy.

This universal (Greenberg 1963) was examined by Ultan (1978) in his 79-language sample and found to hold true⁴.

Greenberg's pioneering work on implicational universals covers the areas of syntax as well as morphology and phonology, and has given rise to more diversified and sophisticated research along this line (e.g. Lehmann 1973, 1978; Vennemann 1974; Comrie 1981; Hawkins 1983, 1994, 1999; Tomlin 1986; Dryer 1988, 1991, 1992).

For his observed implicational correlations between word order properties, Greenberg (1963) attempted to provide an account in terms of 'dominance' (one type as more common than the other) and 'harmony' (one to one correlation of word order types). For instance, his universal 25 "If the pronominal object follows the verb, so does the nominal object" (Greenberg 1963: 72) can be put as: $VO_{Prn} > VO_{Nom}$, which again can be expressed in (5) as a tetrachoric table:⁵

(5)		VO_{Nom}	$ONomV$
	$OPrnV$	X	X
	VO_{Prn}	X	—

From the table, it is clear that VO_{Nom} occurs with either $OPrnV$ or VO_{Prn} down the table and $OPrnV$ occurs with either VO_{Nom} and $ONomV$ across the table, and that $ONomV$ occurs with $OPrnV$ only and VO_{Prn} occurs with VO_{Nom} only. Therefore, the former two (VO_{Nom} and $OPrnV$) are dominant orders, while the latter two ($ONomV$ and VO_{Prn}) are recessive orders harmonic with the order that occurs with them respectively (Greenberg 1963: 76). Greenberg summarised the rule of dominance and harmonic relations as: "A dominant order may always occur, but its opposite, the recessive, occurs only when a harmonic

⁴ Ultan (1978) treats Khasi as an exception to the universals. However, based on some other studies of Khasi (Roberts 1891; Rabel 1961) Eckman et al (1989) do not consider it as a counterexample because there is no inversion in either *Wh* questions or yes/no questions and the *Wh* word's initial position is optional.

⁵ A tetrachoric table is a table of displaying data regarding two parameters on two dimensions, it is very useful for displaying the data of attested and unattested language types in typological studies, and a standard implicational universal can be shown in the table as having three types attested and one type unattested (Croft 1990: 47-55).

construction is likewise present" (1963: 76). In effect, this is the early attempt of explaining word order universals in terms of competing motivations (Croft 1995).

Two other word order typologists Lehmann (1973, 1978) and Vennemann (1974) after Greenberg focused their typological research mainly on harmony. On the basis of Greenberg's (1963) work on basic word order, they reduced his tripartite typology to bipartite OV-VO typology. Lehmann (1973, 1978) formulated *the Fundamental Principle of Placement*, which assumes that verb and object are primary constituents ('primary concomitants') of syntactic construction in the sentence and that the two basic orders – OV and VO – can be used to predict the other orders, i.e. different types of placement of modifiers in relation to each of the concomitants. Similarly, Vennemann (1974) formulated *the Principle of Natural Serialization*, which assumes that in either OV or VO languages word orders tend to be serialised in one direction in terms of 'operators' (modifiers) and 'operands' (heads), that is, either operators before operands or operators after operands. Attractive though this reductionist approach in dealing with word order universals, both Lehmann (1973, 1978) and Vennemann (1974) are critiqued seriously as empirically less adequate than Greenberg's original implicational universals, mainly failing to account for those languages with one or more dominant word orders (Comrie 1981; Hawkins 1983; Croft 1990).

Hawkins (1983), which Croft (1995) regards as the first significant work on word order typology since Greenberg, refined his word order universals on the basis of examining Greenberg's sample as well as his own 336 language sample. In order to make his word order universals exceptionless and quantitatively predicative, Hawkins (1983) formulated more complex implicational universals than Greenberg's. For example, his third revised implicational universal demonstrates a more constrained yet exceptionless multi-valued correlation between various word order properties (Hawkins 1983: 67):

- (6) (III') If a language has Prep and any verb position other than SVO, then if the adjective follows the noun, the genitive follows the noun; i.e. Prep & - SVO > (NA > NG).

Different from Lehmann (1973, 1978) and Vennemann (1974), Hawkins (1983) argued that adposition rather than verb-object order can better predict modifier orders and that his thus-formulated exceptionless universals can account for those counterexamples to the former in a principled way. His *Principle of Cross-Category Harmony* tries to capture the harmonic pattern in terms of operand-operator (head-modifier) relations, that is, operators are most likely to occur on the same side of the operands. More significant is the quantitative dimension in this distributional principle so that the more consistent the cross-category operand-operator relations, the greater the number of exemplifying languages. Hawkins (1983) also proposed two other principles *the Heaviness Serialization Principle* relating to the syntactic size of the modifier (the heavier the modifier, the more rightward positioning it exhibits relative to its head) and *the Mobility Principle* relating to the syntactic size as well as variability of the modifier (the less heavy the modifier, the more likely it moves away from the harmonic pattern). These two principles act as competing motivations for different types of noun-modifier order. While harmony and heaviness principles can account for most of the noun-modifier orders, the mobility principle is used to address a number of exceptions therein.

Different from Hawkins' (1983) multi-valued universals which are based on an unrepresentative sample, Tomlin (1986) made an attempt to build universals relating only to clausal types on a genetically and areally representative sample. On the basis of examining 402 language sampled from a database of 1,063 languages, Tomlin (1986) established the frequency hierarchy of six clausal word orders as:

$$(7) \quad \text{SOV} = \text{SVO} > \text{VSO} > \text{VOS} = \text{OVS} > \text{OSV}.^6$$

⁶ Tomlin (1986) derived the hierarchy by using the referential statistic – Chi-square to determine whether there exists any statistical significance between the actual frequencies of these six types in his sample (SOV – 180, SVO – 168, VSO – 37, VOS – 12, OVS – 5, OSV – 0). In the hierarchy, '=' means 'statistically equal to' and '>' 'statistically more frequent than'. Though the difference between VOS and OVS is significant at the level of 0.05, Tomlin disregards the difference on account of too few languages of both types and uncertainty about the OVS language status.

In order to account for the frequency pattern exemplified in his sample, Tomlin (1986) formulated three functional principles – *the Theme First Principle* (more thematic information precedes less thematic information), *the Animated First Principle* (more animated NP precedes less animated NP) and *the Verb-Object Bonding Principle* (the bondedness of verb and object is tighter than that of verb and subject). Tomlin (1986) argued that the more of the three principles are realised in a basic word order, the greater the number of languages will be found with that order. While the functioning of Tomlin's (1986) three principles seems to correlate well with the frequencies of the basic word orders in his sample, Dryer (1989) calls into question Tomlin's sampling technique which is biased towards SVO, thereby giving rise to SOV = SVO in (7). Moreover, Song (2001: 83-5) is critical of Tomlin's three principles, which he suggests suffer from conceptual disparity.

Another significant step in word order typology is from Dryer's (1988, 1989, 1991, 1992) most extensive and detailed empirical study of word order. What distinguishes Dryer from previous word order typologists is the application of his more reliable and valid sampling technique to a very large database and the examination of a comprehensive range of word order correlations in terms of statistical significance. Dryer (1988) has found that his language sample challenges many of Hawkins's universals regarding the correlation between adjective-noun order and orders of verb-object, adposition-noun, and genitive-noun. For example, in his language sample: "There is no evidence of any relationship between the order of Verb and Object and the order of Adjective and Noun" (Dryer 1988: 191), "There is no clear evidence for the correlation between Adposition-Noun and Adjective-Noun order" (Dryer 1988: 198), and "There is no clear evidence for the correlation between Genitive-Noun and Adjective-Noun order" (Dryer 1988: 200). Based on the evidence drawn from a cross-linguistic database representing 603 languages, Dryer (1991) argues that SVO languages do not behave as 'mixed' a type as have generally been held; rather, they pattern like verb-initial languages in most cases in spite of a few 'mix' characteristics. From another detailed study (Dryer 1992) in which word order correlations between the order of the verb and object with other orders are examined in a sample of 625 languages, he finds that VO and OV are better basic

predictors of word order correlations, hence the return to the VO/OV bipartite typology proposed by Lehmann and Vennemann. Dryer (1992) then argues that *the Head-Dependent Theory*⁷ is inadequate in accounting for these correlations and proposes instead *the Branching Direction Theory* in terms of ordering of phrasal and nonphrasal elements:

(8) *Branching Direction Theory* (BDT):

Verb patterners are nonphrasal (nonbranching, lexical) categories and object patterners are phrasal (branching) categories. That is, a pair of elements X and Y will employ the order XY significantly more often among VO languages than among OV languages if and only if X is a nonphrasal category and Y is a phrasal category.

Based on the distinction between phrasal and nonphrasal categories, Dryer (1992) claims that the Branching Direction Theory predicts the tendency for languages to be consistently left- or right-branching, that is, VO languages tend towards right-branching in which phrasal categories follow nonphrasal categories and OV languages tend towards left-branching in which phrasal categories precede nonphrasal categories. He suggests that the Branching Direction Theory proves to be more effective than the head-dependent theory in that some problematic correlation pairs for the latter (e.g. those of article and noun, and of auxiliary verb and content verb) can now be adequately accounted for. In the sense of Whaley's (1997: 43) distinction between internal explanation ("based on the system of language itself") and external explanation (drawn "on considerations outside of the language system") for typological universals, we can see that the Branching Direction Theory is in effect an explanation from both internal and external perspectives in that it is a claim involving internal syntactic properties which Dryer (1992: 128-32) also relates to processing efficiency.

Since the 1990s, Hawkins has proposed alternative explanations of his word order universals (Hawkins, 1990b, 1993, 1994). Hawkins (1990b) argues that many cross-

⁷ The theory assumes that in linguistic structure, dependents tend to be placed consistently either before or after heads.

linguistic word order patterns follow from simple considerations of ease of parsing in performance, in other words, the human parser in performance strongly influences the competence grammar concerning word-ordering regularities. Instead of the several principles in Hawkins (1983), a new major processing-motivated principle - *the Early Immediate Constituents* is put forth to account for word ordering regularities crosslinguistically (Hawkins 1990b). Hawkins (1993) further illustrates his parsing theory for crosslinguistic regularities of word order on the basis of two principles - the Early Immediate Constituents and *the Mother Node Construction*; the former involves rapid and efficient recognition of syntactic structures while the latter exemplifies such structures in contrast to the heads theory. In effect, this processing approach for addressing crosslinguistic regularities of word order is most profoundly announced in Hawkins (1994), the first book on the systematic investigation of the processing basis for structural regularities both within language and across grammars. This represents a major departure from the earlier head-dependency approach in Hawkins' earlier work.

Incorporating insights from typological studies of language universals, generative grammar and psycholinguistic studies of language processing, Hawkins (1994) postulates a performance theory of order and constituency, which, based on an in-depth analysis of linear ordering in performance and in grammars, claims that "grammars are profoundly shaped by processing" (1994: xi). The Early Immediate Constituents, the major processing mechanism is based on a few additional assumptions such as *the Constituent Recognition Domain* and *the Mother Node Construction*. The Constituent Recognition Domain is defined as (Hawkins 1994: 58-9):

(9) *Constituent Recognition Domain (CRD)*

The CRD for a phrasal mother node M consists of the set of terminal and non-terminal nodes that must be parsed in order to recognize M and all ICs of M, proceeding from the terminal node in the parse string that constructs the first IC on the left, to the terminal node that constructs the last IC on the right, and including all intervening terminal nodes and the non-terminal nodes that they construct.

This can be illustrated by the two sentences in (10).

- (10) a. John VP[wrote NP[the letter that was full of puzzles] PP[to Smith]]
b. John VP[wrote PP[to Smith] NP[the letter that was full of puzzles]]

In both 10a and 10b, the mother node VP has three immediate constituents – V, NP and PP; yet the two VPs differ in one aspect – different constituent recognition domains. The Constituent Recognition Domain for the VP in 10a starts from the first IC, the verb *wrote*, through the lengthy NP in the middle, to the preposition *to* that constructs PP, the last IC in the parse string. The Constituent Recognition Domain for the VP in 10b also starts from the verb *wrote*, but through PP, the two-word IC, to the determiner *the* that signals the construction of NP, the last IC to be parsed. The Constituent Recognition Domain in 10b is shorter than that of 10a in terms of number of terminal and non-terminal nodes or words to be parsed, thereby enhancing a more rapid recognition of the constituent VP. Essential to the Constituent Recognition Domain is the principle of *Mother Node Construction*, which states (Hawkins 1994: 62):

(11) *Mother Node Construction* (MNC)

In the left-to-right parsing of a sentence, if any word of syntactic category C uniquely determines a phrasal mother node M, in accordance with the PS rule of the grammar, then M is immediately constructed over C.

Accordingly, the mother-node-constructing categories uniquely determine the mother node of a constituent. In (10), the V is the mother-node-constructing category of the VP, that is, as soon as the V *wrote* is parsed, it determines the VP as the mother node of the V. The same parsing decision applies to NP and PP in (10) as well. The mother-node-constructing category of the NP is the Det(erminer) and that of the PP is the P(reposition). Once the parser reaches the determiner *the*, the NP is immediately identified as its mother node; in a similar vein, the parsing of the preposition *to* immediately determines the PP as its mother node. The Constituent Recognition Domain is in effect a structural domain where the overall syntactic structure of a constituent, the VP in (10) is determined upon the recognition of all ICs, no matter how many words are

left yet to be processed (one in 10a and six in 10b). In order to quantify relative Constituent Recognition Domain sizes and efficiency levels, Hawkins (1994: 76-7) proposes a metric of calculating ratios of ICs to non-ICs or words in terms of the percentage of the number of ICs divided by the total number of non-ICs or words within a Constituent Recognition Domain.⁸ For example, there are three ICs for the VP Constituent Recognition Domain in both 10a and 10b, but the (italicised) words for that domain are nine and four respectively. As a result, the IC-to-word ratios for the VP Constituent Recognition Domain are different: 3/9 (or 33.3%) in 10a and 3/4 (or 75%) in 10b. The mother node of VP and its ICs are parsed more rapidly and efficiently in 10b than in 10a because the latter has a higher IC-to-word ratio than the former. Based on the above assumptions, Hawkins formulates the principle of *Early Immediate Constituents* (1994: 77):

(12) *Early Immediate Constituents* (EIC)

The human parser prefers linear orders that maximize the IC-to-non-IC ratios of constituent recognition domains.

As the IC-to-word ratio for the VP Constituent Recognition Domain is maximized in 10b (75%) against 10a (33.3%), the Early Immediate Constituents predicts that the former is preferred over the latter for the efficient parsing of the mother node – VP and its immediate constituents – V, NP and PP. The Early Immediate Constituents is applied to test performance data in ten typologically distinct and genetically diverse languages and is claimed to be able to account for the basic word order in Tomlin (1986) and a range of word order correlations in Dryer (1992) as well as word order universals in Hawkins (1983, 1994). Although processing assumption underlies both Dryer's branching direction theory (1992) and the Early Immediate Constituents, the latter seems to be better in making predictions about the ordering of multiple phrasal categories within the

⁸ The IC-to-non-IC ratio for a CRD is usually higher than the IC-to-word ratio because words alone are just the terminal elements excluding the internal non-IC structure standing between ICs and terminal words. Nevertheless, the actual differences between the ratios do not matter as much as the correspondence between the two metrics in terms of relative ranking does, hence the latter used as a shorthand for the former. Furthermore, the IC-to-word ratio is practically advantageous in that it deals with the observable words avoiding varying accounts of the internal non-IC structure, though this metric itself may be problematic in languages with very rich word-internal structure (Hawkins 1994: 74-6, 451).

same syntactic domain apart from the ordering between phrasal and non-phrasal categories, and in making predictions about distributional universals in terms of relative occurring frequencies of different possible orders (Song 2001). The Early Immediate Constituents is proved to provide an effective functional account for both rapid on-line free ordering and the innate grammaticalised ordering of word orders (Hawkins 1994), and "The EIC theory will thus play an important role in future development of word order typology much the same way that Dryer (1992) has formed the solid empirical basis for word-order research" (Song 2001: 110).

2.1.3.2. Grammatical Hierarchies

As has been briefly stated in the previous chapter (1.2), a grammatical hierarchy captures a hierarchical pattern of typological variation in which members of the same grammatical category are ranked relative to one another in a systematic way. For example, in the grammatical category of *number*, the hierarchical ranking for its members is captured in Greenberg's *Universal 34* (1963: 74).

(13) *Universal 34*

No language has a trial number unless it has a dual. No language has a dual unless it has a plural.

Taking into account Greenberg's *Universal 35* (1963: 74) which takes the singular as the basic structurally unmarked number and the non-singular as the marked number, this hierarchy of number can be expressed as:

(14) trial > dual > plural > singular⁹

Recall (from 1.2) that such a grammatical hierarchy can be interpreted as a chain of simple implicational universals or a series of markedness patterns in relative terms. Therefore, the presence of the trial implies the presence of the dual, which again implies the presence of the plural and so on. In this manner, the implicatum of the first universal

⁹ '>' means 'implies' or 'is more marked than'.

becomes the implicans of the second, whose implicatum again becomes the implicans of the third universal. This hierarchical relation is also manifested in the relative values of markedness between the members. For instance, the plural is marked against the singular but unmarked against the dual; similarly, the dual is marked against the plural but unmarked against the trial.

Apart from the hierarchy of number, grammatical hierarchies are discovered to exist in many grammatical aspects of language including hierarchies of grammatical relations, animacy, person, definiteness, NP-type, bondedness and various phonological hierarchies (Croft 1990, 1995). One of the most influential works on hierarchies of grammatical relations is *the Accessibility Hierarchy* of Keenan and Comrie (1977), which has attracted wide-spread attention and invited a huge amount of research in the field of first/second language acquisition as well as in the literature of typology.

Examining the syntactic form of the relative clauses in a sample of about fifty languages, Keenan and Comrie (1977) found that languages vary systematically with respect to relativisation on different NP positions and that certain positions are more accessible to relativisation than others. In order to capture this typological regularity found in relative accessibility to relativisation of NP positions, Keenan and Comrie (1977: 66) formulated the Accessibility Hierarchy.

(15) *Accessibility Hierarchy (AH)*

SU > DO > IO > OBL > GEN > OCOMP¹⁰

To better understand the Accessibility Hierarchy, consider the English examples 16a – 16f which illustrate relativisation on each position of the hierarchy.

- (16) a. The boy *who* got the answer right is clever. (SU)
b. The boy *whom* the teacher praised is proud. (DO)
c. The boy *to whom* the girl lent her car broke his words. (IO)
d. The boy *with whom* the principal talked looks upset. (OBL)

¹⁰ '>' means 'is more accessible than'; SU = subject, DO = direct object, IO = indirect object, OBL = oblique object (or major oblique case NP expressing arguments of the main predicate), GEN = genitive (or possessor) NP, and OCOMP = object of comparison (Keenan and Comrie 1977: 66).

- e. The boy *whose* father teaches law stole the jewels. (GEN)
- f. The boy *who* you run as fast as won the race. (OCOMP)

As illustrated in (16), all the positions on the hierarchy in English can be relativised on; however, the majority of the world's languages behave differently from English with different types of relativising possibilities. Nonetheless, Keenan and Comrie (1977: 67) suggested that all languages conform to the constraints of the Accessibility Hierarchy in (17).

(17) *The Hierarchy Constraints (HCs)*

- a. A language must be able to relativize subjects.
- b. Any RC-forming strategy must apply to a continuous segment of the AH.
- c. Strategies that apply at one point of the AH may in principle cease to apply at any lower point.

In accordance with the Hierarchy Constraints, all languages must allow relativisation on the subject, whether the relativisation on other positions is allowed or not; therefore, a relativisation strategy on subjects in a language is justifiably taken to be 'a primary strategy' in that language. Accordingly, if a primary strategy can apply to any position on the hierarchy, then all the higher positions can be relativised on by the same strategy in a contiguous manner. For instance, if a primary strategy can apply to IO in a language, it must be able to relativise on DO and SU in that language as well, and cannot simply 'skip' DO and go back on SU again. This constraint (Hierarchy Constraint b) also applies to non-primary RC-forming strategies such as pronoun-retention or obliteration (gapping) strategy in a restricting clause. A primary strategy may stop at any cut-off position on the Accessibility Hierarchy in a language, thereby illegitimising its application to the lower positions. In effect the data provided in Keenan and Comrie (1977, 1979) lend support to the theoretical proposals of the Accessibility Hierarchy and the Hierarchy Constraints.

In an attempt to explain the hierarchy constraints of the Accessibility Hierarchy, Keenan and Comrie proposed a simple processing account in terms of 'psychological ease of comprehension' (1977: 88), with which relative clauses formed on higher positions on

the Accessibility Hierarchy are easier to understand than those formed on lower positions. The psychological validity of the Accessibility Hierarchy is speculatively suggested as based on 'a recognition strategy' and 'independent reference' (Keenan and Comrie 1977: 93-5). The former indicates a likely coincidence of the NP positions on the Accessibility Hierarchy with those NPs that are found in simple sentences in terms of frequencies of occurrence. The latter refers to the relative accessibility of the subject over other NPs as being due to the inherent nature of independent reference of the subject the other NPs lack; as a result, relativisation on a non-subject produce a structure with two 'necessarily independently referring expressions' (instead of one on a subject), hence psychologically more difficult.

The emergence of the Accessibility Hierarchy has stimulated a lot of research within the field in an attempt to test its validity. Based on a sample of 105 languages, Comrie (1981) observed some regularity in the application of different relativisation strategies across the Accessibility Hierarchy, that is, within individual languages, more explicit relativisation strategies are used for the lower NP positions down the Accessibility Hierarchy, while less explicit relativisation strategies are used for the higher NP positions up the hierarchy.¹¹ He went on to propose a functional explanation for this generalisation: "the more difficult a position is to relativize, the more explicit indication is given of what position is being relativized, to facilitate recovery of this information" (Comrie 1981: 156).

Lehmann (1986) examined the correlations between properties related to relativisation on the Accessibility Hierarchy. He (1986: 672) found that relative clauses may be nominalised to varying degrees and that the degree of RC nominalisation correlates with the positional types of RCs (prenominal and postnominal) and with the degree of accessibility to the NP positions on the Accessibility Hierarchy. Prenominal RCs are found to have higher degree of nominalisation and lower degree of accessibility to the NP

¹¹ In terms of the explicitness of relativisation strategies, the obliteration strategy (which does not encode the head noun) is less explicit than the relative-pronoun strategy (which encodes the head noun as a pronominal word maybe with case, e.g. *that*, *who*, *whom* and *whose*), which in turn is less explicit than the

positions on the Accessibility Hierarchy than postnominal RCs. Furthermore, the degree of nominalisation and the positional type in combination can explain why pronoun-retention and relative-pronoun strategies are found in postnominal rather than prenominal RCs. Both personal pronouns and relative pronouns are anaphors indicating coreference between the head and the relativised position, hence the greater occurrence of the two strategies in the more sentential, less nominalised postnominal RCs (Lehmann 1986: 674). The very nature of anaphoric expressions that they are assumed to be placed after a referent (e.g. a head noun), may also contribute to the preference of the two strategies in the postnominal RC type (Lehmann 1986: 676).

In terms of argument coding, the nominative-accusative relation is most common pattern in the languages of the world, while the ergative-absolutive relation is the other very common pattern (Comrie 1978).¹² With respect to accessibility to relativisation in ergative-absolutive languages, it is shown that the absolutive relations (S and P) are the only accessible positions in the Mayan languages (Larsen and Norman 1979) and Dyirbal (Dixon 1979). Fox (1987) reevaluated the "subject primacy" of the Accessibility Hierarchy and argued that the role of subject does not have a special cognitive prominence. Through a statistical analysis of relative clauses obtained from telephone and face-to-face conversations in English, Fox found that even in nominative-accusative-oriented languages such as English, the discourse preferences in relativisation favour the absolutive – the relativisation of intransitive subject or object rather than transitive subject (82 instances of the former versus 10 instances of the latter; Fox 1987: 858). To modify the Accessibility Hierarchy, she proposed *the Absolutive Hypothesis*, which holds that "every language which has a strategy for relativizing must be able to relativize on at least S and P" (Fox 1987: 864). Assuming that the category absolutive rather than subject occupies the leftmost position on the Accessibility Hierarchy, this hypothesis is found to make correct predictions about relativisation phenomena in Dyirbal and Mayan

pronoun-retention strategy (which also encodes the head noun as a pronominal word, but maybe with case, gender and/or number, e.g. *he, him, she, them* and *it*).

¹² In nominative-accusative languages, the intransitive subject S is coded the same way as the transitive subject A (agent), hence both referred to as the 'nominative' (traditionally as the 'subject'), while the transitive object P (patient) is referred to as the 'accusative' (traditionally as the 'object'). On the other

languages. However, whether this hypothesis is universally valid remains to be further tested crosslinguistically (Whaley 1997: 264; Song 2001: 241).

Relative clause types (prenominal/RelN and postnominal/NRel) are shown to correlate with word order types (VSO, SVO and SOV) in Greenberg's (1963) data for illustration of his Universal 24, which is presented in Table 2.1.

Table 2.1.

	VSO	SVO	SOV
RelN	0	0	7
NRel	6	12	2
Both	0	1	1

(Adapted from Greenberg 1963: 71)

Although Greenberg did not directly expound the correlation between RC types and word order types, it is clear from Table 2.1, that SVO, like VSO, correlates significantly with NRel, while SOV correlates predominantly with RelN. Relationships between RC types and word order types have subsequently been examined more extensively by Mallinson and Blake (1981), Keenan (1985) and Dryer (1991). Apart from many other findings, all these studies have confirmed what is embodied in Greenberg's data, that is, verb-medial languages behave more like verb-initial languages rather than verb-final languages with respect to RC types.

In seeking explanation for the distribution of RC types and the phenomena of relativisation concerning the Accessibility Hierarchy, processing, again, plays an important role. Kuno (1974) argued that the correlation between NRel and VSO, and between RelN and SOV can be explained in terms of processing constraint by which the human parser bound by the human capacity of short-term memory tends to avoid the centre-/self-embedded structure. Because centre-embedding creates processing difficulties, the above two types of correlation are ideal in that they reduce such processing difficulties to minimum by allowing only one centre-embedded restricting

hand, in ergative-absolutive languages, S is coded the same way as P, hence both referred to as the 'absolutive', while A is referred to as the 'ergative' (Comrie 1978).

clause in each. On the other hand, VSO with RelN and SOV with NRel will each have three instances of centre-embedding (i.e. centre-embedded restricting clauses before three arguments - S, A, P in VSO, and after S, A, P in SOV), thereby maximising processing difficulties in these two types of permutations.

Mallinson and Blake (1981: 300-1) regarded Kuno's processing explanation for the two permutations - NRel in VSO and RelN in SOV - as particularly appealing. Extending Kuno's account to V-medial languages which have a maximum of two centre-embedded RCs, they (1981: 310) claimed that:

Verb-medial languages present a degree of self-embedding whichever order of Head/RC is adopted, and the choice is therefore less clearcut than in verb-initial and verb-final types. The fact that verb-medial languages tend to have Head-RC order apparently supports the view that Head-RC order is the unmarked order and is only abandoned if there is overwhelming pressure to take RC-Head order, as is the case in verb-final languages like Japanese.

As has been shown in 2.1.3.1, the role of processing underlies Dryer's (1992) Branching Direction Theory and Hawkins' (1994) Early Immediate Constituents, both of which can account for most of the observed correlations between the basic word order types and other types of ordering including the order of N (head)/RC. However, Hawkins (1994) takes a step further in addressing the phenomenon unaccountable by the above theories, namely, the evidence of both RelN and NRel orders in V-final languages. Hawkins (1994: 324) argues that in order to avoid "a structural misanalysis and garden path", the Early Immediate Constituents may give way to "immediate matrix disambiguation" in which "the parser makes an immediate decision about the main- or subordinate-clause status of a clause". As a result, RelN in V-final languages predicted by the Early Immediate Constituents may be rearranged by the principle of immediate matrix disambiguation to become NRel in V-final languages instead.

Hawkins' (1994: 37-46) processing explanation is also exemplified in his account for the Accessibility Hierarchy. He (1994: 38) proposes "a complexity metric" in support of Keenan and Comrie's claim "The AH directly reflects the psychological ease of comprehension". The metric, *Minimal Structural Domains*, which can be used to quantify the structural complexity associated with the positions in the sentence in a consistent way, is defined by Hawkins (1994: 39) in (18):

- (18) The Minimal SD of a node X in C consists of the smallest set of structurally integrating nodes that are grammatically required to co-occur with X in C.¹³

For a language with an English-type syntax, for example, the minimal structural domain of the subject (which requires only two nodes) is less complex than that of the direct object (which requires four nodes), which, in turn, is less complex than that of the oblique/indirect subject (which requires six nodes), and so on (Hawkins 1994: 41). Similarly, the complexity of the relativisation domains also increases as the relativised position goes from the SU position down the Accessibility Hierarchy, because the relativisation domains for all positions on the hierarchy, in addition to their respective minimal structural domains described above, contain the same higher structure – the dominating NP and S-bar nodes and the head noun (Hawkins 1994: 31). As structural domains increase in complexity down the hierarchy, the processing load increases and these performance data correlate closely with the conventionalised grammatical rules of particular grammars, hence the conformity to the Accessibility Hierarchy across languages (Hawkins 1994: 30). In this way, crosslinguistic generalisations such as the Accessibility Hierarchy can be consistently accounted for in terms of processing-based structural complexity, which is a step forward in processing account for typological universals in terms of internal and external explanation.

¹³ The structural domain of a node X in a tree dominated by a constituent C consists of all the nodes that structurally integrate X in C, including all nodes dominating X within C (including C itself), all or some sisters of X, and all sisters of the nodes dominating X within C; the structural complexity of a structural domain is measured by calculating the set of nodes within the domain (Hawkins 1994: 25-30).

Recently, this processing approach of Hawkins' (1999) is furthered in his examination of cross-linguistic variation concerning principally *wh*-questions and relative clauses in terms of filler-gap dependencies. The processing of filler-gap dependencies is explicitly expounded in his model of language processing, for he believes (Hawkins 1999: 245):

that the conventions themselves have been profoundly shaped by processing, and that by adding processing as a central component to our grammatical theory we can build a richer theory: better predictions can be made for grammatical variation and for possible versus impossible grammars; and a lot of what is currently stipulated can be motivated and derived.

Essential to this approach are the basic concepts of *filler*, *gap*, and *filler-gap domain*. *Filler* refers to the moved element and *gap* refers to its trace, an empty position which is either subcategorised or nonsubcategorised; a filler in surface structure must be matched with its gap or only with its subcategoriser if there is one (Hawkins 1999: 244-6). Consider the examples in (19):

- (19) a. *Who_i [do you think that the lecturer taught O_i] ?*
b. *The lecturer_i [that O_i teaches] these students is from Japan.*
c. *The students_i [that the lecturer teaches O_i] are well motivated.*
d. *The student_i [that the lecturer gave books to O_i] is my brother.*

The filler in a *wh*-question such as 19a is the *wh*-word *who*, which is matched with a co-indexed gap after its subcategoriser *taught*. The filler in a relative clause is the head noun such as *the lecturer*, *the students*, and *the student* in 19b, 19c, and 19d respectively, while the relativiser *that*¹⁴ is "a FILLER COPY within the relative clause, which, by being fronted to a position adjacent to the head, leaves a gap in the position relativized on, just as the deletion strategy leaves a gap" (Hawkins 1999: 251). The filler *the lecturer* is matched with a co-indexed gap after *that* in 19b, so is *the students* after *teaches* in 19c and *the student* after *to* in 19d. Structures containing filler-gap dependencies vary greatly in these

¹⁴ Particles or words which introduce relative clauses are called "relativizers" or "relative pronouns"; "In English, for example, *that* is a relativizer, whereas *who*, *whom*, *which*, and *whose* are relative pronouns" (Whaley 1997: 259).

sentences; in general these structures are difficult to process in terms of processing complexity in identifying the correct gap with its filler. In other words, before a gap can be correctly identified and ultimately filled with its filler, the filler has to be held in working memory, to be constantly processed against all the material on the path to the gap site, and finally to be released from working memory by locating its correct gap site. Hawkins defines the domain containing a filler-gap dependency as:

(20) *Filler-Gap Domain (FGD):*

An FGD consists of the smallest set of terminal and nonterminal nodes dominated by the mother of a filler and on a connected path that must be accessed for gap identification and processing; for subcategorized gaps the path connects the filler to the gap's subcategorizer and includes, or is extended to include, the gap's dependent and disambiguating arguments (if any); for nonsubcategorized gaps the path connects the filler to the gap site; all constituency relations and cooccurrence requirements holding between these nodes belong to the description of the FGD.

(Hawkins 1999: 248)

Hawkins (1999: 251) explains that the complexity of filler-gap domains varies in proportion to their size and in proportion to the amount of simultaneous syntactic and semantic processing that is required. With respect to processing load, the human parser tends to minimise filler-gap domains if possible and such filler-gap domain minimisation can be achieved through reducing the set of nodes in the domain. In terms of the set of nodes to be processed in the filler-gap domain, the filler-gap domain of 19b is smaller than that of 19c, which, in turn, is smaller than that of 19d. The Accessibility Hierarchy of Keenan and Comrie (1977), which has been explained by Hawkins (1994) as involving increasing complexity for the relativisation positions down the hierarchy in terms of the number of nodes and structural relations that need to be counted, is revisited by Hawkins (1999) in terms of filler-gap domains. The Accessibility Hierarchy is presented in (21) (Hawkins 1999: 253):

(21) *Accessibility Hierarchy (AH):* SU > DO > IO/OBL > GEN

The minimal filler-gap domain for each relativisable position on the hierarchy includes the filler N, the subcategoriser of N's gap, the gap's overt dependent arguments, and all nodes dominated by the mother of N (NP) that are required for grammaticality. Hawkins (1999: 255) sums up the set of nodes to be processed in the minimal filler-gap domain for the positions on the Accessibility Hierarchy as follows:

(22) Minimal FGDs for relativizations on:¹⁵

SU = 5	{N, NP, V, VP, S}
DO = 7	{N, NP, N _s , NP _s , S, V, VP} (requires SU)
IO = 9	{N, NP, N _s , NP _s , S, V, VP, N _o , NP _o } (requires SU & DO)
OBL = 9	GEN-SU = 9 GEN-DO = 11 GEN-IO = 13 GEN-OBL = 13

The processing complexity associated with 19b, 19c and 19d can thus be explained in terms of the relative size of the filler-gap domain each of the three has. That is, the filler-gap domain of 19b (containing 5 nodes for processing) is more accessible than that of 19c (containing 7 nodes), which again is more accessible than that of 19d (containing 9 nodes). This processing approach provides a possible explanation for typological universals. Not only can the processing approach naturally explain internal and external factors that underlie some universals, but many more descriptive insights and implicational patterns have been motivated in Hawkins (1999), some of which are reviewed in 2.1.3.3, and tested against interlanguage data in the present study. In addition, "this approach provides an alternative to the assumption of innate parameterized subjacency constraints in this area" (Hawkins 1999: 244).

More recently, Hawkins (2001) again argues that a common principle of processing efficiency explains the patterns of adjacency both in performance data and in conventionalized adjacency preferences of these performance data across languages. In a word, Hawkins' (1994, 1999) processing approach not only provides an illuminating

¹⁵ For the sake of illustration and space, the set nodes in the minimal FGDs for relativisation on lower than IO are not cited. For a detailed demonstration, refer to Hawkins (1999: 253-5).

account for the performance data and crosslinguistic grammatical conventions regarding typological universals and hierarchies, it may prove equally significant in addressing the performance data and interlanguage features in SLA as well.

2.1.3.3. Some Recently Motivated Implicational Hierarchies/Hypotheses

As has been discussed above (2.1.3.2), the extent of accessibility to each relativised position on the Accessibility Hierarchy depends on the structural complexity of the filler-gap domain for that relativised position and the simultaneous syntactic processing required for gap identification across that filler-gap domain (Hawkins 1999). The higher the relativised position (e.g. SU) is on the Accessibility Hierarchy, the smaller its size of filler-gap domain will be and the less its syntactic processing for gap identification is required, hence the more accessible that position to relativisation. This processing perspective of Hawkins' (1999) has not only accounted for the Accessibility Hierarchy more profoundly than previously attempted but also motivated a number of implicational universals/hypotheses regarding grammatical variation across grammars, some of which are addressed below.

One of the significant implicational formulations of Hawkins (1999) is the *Clause Embedding Hierarchy* for filler-gap domains, which provides an alternative account of general constraints of grammar to both the *Complex NP Constraint* of Ross (1967) and the *parameterised subjacency* of Chomsky (1981b, 1986).¹⁶ In explanation of the constraints of grammar imposed on movements out of complex structures, the Complex NP Constraint (Ross 1967) allows no element contained in an S dominated by an NP with a lexical head noun to be moved out of that NP by a transformation, and, subsequently, Chomsky's theory of subjacency (1981b, 1986) disallows elements to be moved out of all environments that contain two or more bounding nodes. On the basis of the examination of the two assumptions and relevant data across grammars, Hawkins finds that "general constraints of grammar have been proposed that block filler-gap dependencies in certain environments with large FGDs, while smaller FGDs may be unconstrained and allowed to run free" (1999: 262).

¹⁶ The italicised bold letters are used in this section for those universals tested in this study.

By analysing the languages which have imposed grammatical constraints on filler-gap domains differently, Hawkins concludes that: "It appears that infinitival phrases are most hospitable to gaps, that finite subordinate clauses are more resistant, while complex NP environments are most resistant of all", and formulates a hierarchy for gaps in clause embeddings as follows (1999: 263):

(23) *Clause Embedding Hierarchy:*

$$s[\dots VP[\dots O_i \dots] \dots] > s[\dots S'[\dots O_i \dots] \dots] > s[\dots NP[Det N S'[\dots O_i \dots] \dots]]$$

In the hierarchy, S stands for the clause that is immediately adjacent to the filler, VP' for an embedded infinitival phrase, S' for a finite subordinate clause, and NP-S' for a complex noun phrase comprising a dominating noun phrase over a finite subordinate clause. This hierarchy can adequately address filler-gap dependencies in clause embeddings cross-linguistically: different languages have different cut-off positions of grammaticality on the hierarchy with all those above the cut-off position regarded as grammatical. Moreover, it can not only account for the violations of the Complex NP Constraints/subjacency found in some languages, but can avoid some theory-internal controversies of the subjacency theory such as the successive cyclic movement of *wh* (Hawkins 1999: 262-6).

This hierarchy is, in essence, motivated by the differences in the sizes of the filler-gap domains of the respective positions of the hierarchy. The lower the position down the hierarchy, the larger the size of its filler-gap domain and the more nodes the human parser has to process. Consider the following examples of the relevant structures in English cited from Hawkins (1999: 264):

- (24) a. The person_i s[that you tried VP[to see O_i]] is Harry.
 b. The person_i s[that you hoped s[that you would see O_i]] is Harry.
 c. *The student_i s[who_i you know NP[the professor s[that taught O_i]]] is Harry.

The filler-gap domain of 24a proceeds from the filler *the person* to its gap's subcategoriser *see*, and includes a higher verb *try* with its subject *you*, and an embedded

infinitival phrase. The filler-gap domain of 24b is more complex and contains more nodes, since the subcategoriser of the gap now occurs within a finite subordinate clause containing an additional overt subject *you* and finiteness-marked *would* (albeit the infinitival complementiser *to* in 24a is comparable to the complementiser *that*). As compared with 24b, 24c is even more complex in that there is an additional dominating NP *the professor* over S' and the additional head noun *professor* also c-commands the gap. From 24a to 24c, there are therefore more nodes to be processed and "these additional nodes involve phonological, morphological, syntactic and semantic processing operations that apply simultaneously with filler-gap processing... Ultimately, of course, it is the aggregate processing load associated with these positions in performance that leads to their ranking and to the cut-off points in grammatical conventions" (Hawkins, 1999: 264). As shown by the examples, English cuts off at the complex NP position.

For English, permitted filler-gap domains for both relative clause heads and *wh*-question fillers cut off at the complex NP position; fillers are permitted to be matched by gaps in infinitival phrases and finite subordinate clause. However, in the structure with a *wh*-question filler, if the subordinate clause S' contains another gap apart from the gap to be matched by the *wh*-question filler, then this type of structure cuts off at the finite complement rather than the complex NP position on the hierarchy, as in the following examples cited from Hawkins (1999: 269):

- (25) a. What_i did you wonder [how_j to bake O_i O_j]?
 b. *What_i did you wonder [how_j they would bake O_i O_j]?

Similar to the case of (24), 25a with an infinitival complement has a smaller size of filler-gap domain, hence less nodes to be processed, than 25b with a finite subordinate clause. In fact, the *wh*-island construction (Chomsky 1973) in 25b has no *wh*-island effect in 25a that contains an indirect question infinitival.

As has been demonstrated, it is the relative size of a filler-gap domain that determines the relative structural complexity and the processing load associated with it. From this point, Hawkins goes on to argue for a more general minimisation hypothesis for syntactic

representations in filler-gap domains of the same size which may exhibit subtle degrees of preference. This hypothesis is presented as follows (Hawkins 1999: 269):

- (26) *Reduce Additional Syntactic Processing in FGDs*: The human processor prefers to minimize the syntactic rules and processing operations that apply in FGDs.

Based on this hypothesis and the supportive evidence across grammars, Hawkins formulates an implicational prediction – the *Subordinate Gap/No Gap Hierarchy* (Hawkins, 1999: 269):

- (27) *Subordinate Gap/No Gap Hierarchy*: If a matrix filler can be matched with a gap in a subordinate clause of complexity n containing another gap, then it can be matched with a gap in subordinate clause of complexity n containing no other gap.

Consider the following two examples of Hawkins (1999: 269):

- (28) a. What did you hope [that they would bake O_i]?
b. *What_i did you wonder [how_j they would bake O_i O_j]?

The surface filler-gap domains of 28a and 28b appear to contain the same quantity of nodes for processing. However, the differences between the syntactic rules that apply in 28a and 28b respectively make the former grammatical and the latter ungrammatical in English. That is, as compared with 28a in which the subordinate clause contains only one gap to be matched with the matrix filler, in 28b (a *wh*-island construction), the subordinate clause contains another gap apart from the gap to be matched with the matrix filler, thus more difficult for processing.

This hierarchy also applies within the complex NP constructions. Consider the following examples (Hawkins, 1999: 270):

- (29) a. *Who_i s[do you know NP[the professor_j s'[that O_i taught O_i]]]?
b. *What_i s[do you regret NP[the fact s'[that he stole O_i]]]?

Even both 29a and 29b are regarded as ungrammatical in English. Movements out of NP complements such as 29b should be grammatically better than out of relative clauses in 29a because 29a contains a second gap co-indexed with the head of the complex NP but 29b does not.

Apart from the human processor's preference for reducing additional syntactic processing within a filler-gap domain, Hawkins (1999) also argues for its preference for less semantic processing within the domain. Hawkins (1999: 271) summarises a similar minimisation hypothesis for semantic representations in filler-gap domains as:

- (30) *Reduce FGDs semantically*: the human processor prefers to minimize the amount of semantic information that needs to be processed in an FGD.

In review of some relevant data in the literature (Kluender 1992 and others), Hawkins suggests that this hypothesis of semantic minimisation motivates two implicational hierarchies – the Bridging Verb Hierarchy and the Head Noun Phrase Hierarchy (Hawkins 1999: 271-3), which share the similar semantic/processing underpinnings as the Hierarchy for Complementisers (Kluender 1992).

Kluender (1992; cited in Hawkins 1999: 273) proposes the *Hierarchy for Complementisers* as follows (where > means 'better than'):

- (31) What did John doubt (a) *that* she would win? >
(b) *if* she would win? >
(c) *whether* she would win?

Kluender provides justification for the hierarchy as: "The complementizer *that* merely signals that a proposition follows, while *if* indexes a possible state of affairs from among an infinite set of such possible states, and *whether* indexes one of only two possible (alternative) states of affairs, and can thus in a way be said to be more referentially specific in character" (1992: 240). Therefore, within the filler-gap domain of (31) starting from the filler *What* and ending with its subcategorisor *win*, 31a (with the least semantically loaded *that*) is better than 31b (with *if*), which, in turn, is better than 31c (with the most semantically loaded *whether* amongst the three complementisers). Or put

in another way, grammaticality declines down the hierarchy as the processing load associated with the complementisers within the filler-gap domain increases.

Based on the previous work (Kluender 1992 and others), Hawkins (1999: 272) formulates the *Bridging Verb Hierarchy*, which states:

- (32) *Bridging Verb Hierarchy*: If a bridging verb or verb complex V of semantic specificity n is grammatical in an FGD, then all verbs or verb complexes V' with less semantic specificity than V will also be grammatical.

Hawkins (1999: 272) cites the following examples from the previous work (Kluender 1992 and others) to justify this hierarchy:

- (33) a. How angry did Mary *say* that John was? >
b. How angry did Mary *say softly* that John was? =
c. How angry did Mary *whisper* that John was?

In (33), *say* is a semantically weak verb, which has less semantic specificity than *whisper* (equal to *say softly*). "Degrees of specificity could be defined in terms of semantic components or features. *Whisper* combines components of meaning associated with both *say* and *softly*, and is accordingly semantically richer and more specific. Specificity could also be defined in terms of entailment: *whisper* entails *say*, but not vice versa" (Hawkins 1999: 272). Within the filler-gap domain (from the filler *How angry* to its gap site after *was*) in (33), 33a with a less semantically specified bridging verb *say*, which readily permits a gap in a *that*-clause, is therefore grammatically better than 33b and 33c both with more semantically specified verbs.

In his *Head Noun Phrase Hierarchy*, Hawkins (1999: 272) states:

- (34) *Head Noun Phrase Hierarchy*: If an NP with head noun (phrase) N of semantic specificity n on the path from filler to gap is grammatical, then all head noun (phrases) N' with less semantic specificity than N will also be grammatical.

Again, Hawkins (1999: 272) justifies this hierarchy by citing examples from the previous work (Kluender 1992 and others) as evidence, which are presented in (35):

- (35) a. Who did you see *a* picture of? >
 b. Who did you see *the* picture of? >
 c. Who did you see *John's* picture of?

Hawkins justifies this hierarchy as follows: "A definite NP adds a uniqueness claim to the existential claim of the indefinite, as well as pragmatic appropriateness conditions guaranteeing the satisfaction of uniqueness within some portion of the universe of discourse (Hawkins 1991). A possessive modifier contains an additional referring expression and defines a relation of possession or association between this referent and the head" (1999: 272). As a result, with the filler-gap domain of (35) beginning with the filler *Who* and ending with its subcategoriser *of*, 35a (with the least semantically specified indefinite article) is grammatically better than 35b (with more semantically specified definite article), which, in turn, is grammatically better than 35c (with the most semantically specified possessive modifier). In other words, as the amount of semantic processing load carried by the head noun in an *of*-complement increases from 35a to 35c, the grammaticality of a filler-gap structure declines simultaneously.

Finally, Hawkins (1999: 278) formulates a hypothesis – *Valency Completeness* which states:

- (36) *Valency Completeness*: The human processor prefers FGDs to include the subcategorizers for all phrases within the domain that contain the gap.

Hawkins (1999: 278) claims that "valency completeness facilitates processing by making it clear how all the phrases containing the gap are structurally connected within the FGD" and that the processing-motivated hypothesis can account for the Sentential Subject Constraint (Ross 1967) and other types of subject-object asymmetry in English. The examples Hawkins (1999: 277-8) uses to justify this hypothesis are presented in (37).

- (37) a. *Who_i s[_idid s'[that Mary disliked O_i] vp[surprise Sue]]?
 b. Who_i s[_idid it vp[surprise Sue s'[that Mary disliked O_i]]]?

- c. *What_i s[_idid NP[the title of O_i] VP[amuse John]]?
- d. What_i s[_idid John VP[read NP[the title of O_i]]]?
- e. *What_i s[_idid VP[_ito read O_i] VP[fascinate Sue]]?
- f. What_i s[_idid it VP[fascinate Sue VP[_ito read O_i]]]?

Among the three pairs of subject-object asymmetries in (37), thirty-seven a, 37c and 37e all have a filler extracted out of the subject component (sentential subject, prepositional complement and infinitival complement respectively). Thus the subcategorisers (*surprise*, *amuse* and *fascinate* respectively) for all the complements that contain the gap stand outside the filler-gap domain resulting in valency incompleteness within the domain, hindered processing and ungrammaticality of the sentences. In contrast, thirty-seven d has a filler extracted out of the object component and both 37b and 37f have an extraposition structure, but the subcategorisers (*surprise*, *read* and *fascinate*) for the containing complements in 37b, 37d and 37f respectively stand within the filler-gap domain, hence valency complete.

In closing, Hawkins' (1999) processing-motivated implicational hierarchies/hypotheses are both theoretically rigorous and empirically attested across grammars, some of which have been demonstrated in this section and will be tested against interlanguage data of the study.

2.1.4. Summary

While the traditional typological studies focus on morphological classification in an attempt to find differences among languages, modern linguistic typology has focused on research that involves studies of language universals as well as typological investigation. Apart from typological classification of languages, typological comparison is used as a powerful tool in the latter for finding what languages have in common, i.e. typological universals. In fact the universals research of modern linguistic typology has revealed more than the classical morphological typology about the crosslinguistic patterns in particular and the nature of language in general. In view of what has been reviewed above, modern linguistic typology is characterised as:

- 1) In general, the study of languages is both empirically and theoretically oriented, albeit with varying emphasis among different approaches. That is, modern linguistic typology relies heavily on an increasing number of a wide range of languages of the world, which in turn provide evidence for theoretical insights. For instance, crosslinguistic generalisations are formulated on the basis of examining empirical data from large language samples (e.g. Greenberg 1963; Hawkins 1983; Dryer 1992). In order to account for crosslinguistic patterns and constraints on language, typologists have also been seeking internal explanations within the language system itself and/or external explanations outside the language system relating to discourse, processing, economy, perception-cognition, and iconicity (Whaley 1997).
- 2) With its diversity of theoretical orientations and practical methodologies, modern linguistic typology, particularly the Greenbergian typology has become 'a subdiscipline of linguistics' or a well-articulated 'approach' (Croft 1990). Even in terms of its magnitude of research publications, its numerous research centres and institutes, and its crossdisciplinary influence, it should be regarded as a well-established field now. The field promises a substantial development with more languages of the world to be studied and more descriptive and theoretical insights to be gained, and the typological studies of languages may eventually help to answer the question "what is a possible human language?".
- 3) Finally, no less significant is the extension of the approach and findings of modern linguistic typology to studies in the fields of first and second language acquisition. For example, the Accessibility Hierarchy of Keenan and Comrie (1977) has stimulated a huge amount of research in both fields and the present study is another attempt to test the validity and predictive power of this universal as well as some others (Greenberg 1963; Kluender 1992; and Hawkins 1999).

2.2. Second Language Acquisition

This section includes an overview of the past and current theories and practices in the field of second language acquisition (2.2.1) and a more detailed account of the acquisition of English relative clauses by L2 learners from a wide range of L1 backgrounds with particular reference to the Accessibility Hierarchy (Keenan and Comrie 1977) (2.2.2).

2.2.1. An Overview: Theory and Practice in Second Language Acquisition

Second language acquisition (SLA) research can be regarded as having been established as a field of inquiry in applied linguistics around the late 1960's (Ellis 1994). Though with a fairly short history, there is no lack of empirical studies and theoretical discussions in this field. The different approaches associated with these different theoretical and empirical orientations are briefly addressed hereinafter.

2.2.1.1. Contrastive Analysis

During the 1960's, SLA research was predominantly conducted from the perspective of *Contrastive Analysis*, which was based on behaviorist psychology and structuralist linguistics. The Contrastive Analysis sought to explain SLA by predicting the errors that learners make through identifying the linguistic differences between their first language and the target language. The Contrastive Analysis approach argued that on the one hand, learners' errors occurred primarily as a result of interference when they transferred L1 grammatical 'habits' into L2, and interference occurred whenever grammatical 'habits' of L1 and L2 differed; and on the other hand, by systematically comparing language habits of L1 and L2, errors could be predicted, avoided and even overcome through practice and new habit formation in L2 learning.

For example, Brooks (1964) worked out a grammatical circle of pattern practice, insisting that in SLA, structure drill is essential and should be furnished by analogy as well as by analysis. The behaviourist underpinning of the Contrastive Analysis is clearly spelt out in Brooks (1964: 58):

Like sin, error is to be avoided and its influence overcome... The principal method of avoiding error in language learning is to observe and practice the right model a sufficient number of times; the principal way of overcoming it is to shorten the time lapse between the incorrect response and the presentation once more of the correct model.

Similarly, Lado (1983), one of the leading proponents of the Contrastive Analysis, devised both general and specific procedures for comparing two grammatical structures, and argued that "The list of problems resulting from the comparison of the foreign language with the native language will be the most significant list for teaching, testing, research, and understanding" (Lado 1983: 19).

However, the Contrastive Analysis approach has been found to be inadequate alone as an account for developmental patterns in SLA. That is, the approach can not account for those non-target features of the learning processes through which learners actually perceive, organise and use the target language, hence giving rise to the emergence of a new approach known as the Error Analysis and Interlanguage Theory.

2.2.1.2. Error Analysis and Interlanguage Theory

Around the late 1960's and early 1970's, a number of empirical studies (e.g. Dulay and Burt 1973) showed that many of the errors which were not predicted by the Contrastive Analysis did occur, thereby questioning the validity of the Contrastive Analysis in SLA. Not only was the importance of L1 interference questioned, but the theoretical basis for the Contrastive Analysis---behaviorism was seriously attacked by the mentalist theory (e.g. Chomsky 1959) as well. As a result, the Contrastive Analysis fell into disfavour in SLA and to replace its dominance, a new approach – the Error Analysis (Corder 1967) and the interlanguage theory (Selinker 1972) came into being. The focus for research was consequently shifted from input to output, i.e., from the comparison between L1 and L2 systems to the analysis of learner's interlanguage.

Assuming that both the adult learning a second language and the child learning a primary language may resort to the same innate mechanism, Corder (1967: 164) argued that "we may also postulate that the procedures or strategies adopted by the learner of the second language are fundamentally the same". In this mentalistic view, the learner's errors, like the child's non-standard forms in early stages of L1 acquisition, could provide evidence for the development of the system of a "built-in syllabus" (Corder 1967: 166) for both the child and the adult, and these errors were themselves systematic. In contrast to the Contrastive Analysis, which treated the learner's errors as undesirable, the Error Analysis viewed errors as providing evidence of the system of the language the learner is using. Corder summarised that the learner's errors were significant in three different ways:

First to the teacher, in that they tell him, if he undertakes a systematic analysis, how far towards the goal the learner has progressed and, consequently, what remains for him to learn. Second, they provide to the researcher evidence of how language is learned or acquired, what strategies or procedures the learner is employing in his discovery of the language. Thirdly, (and in a sense this is their most important aspect) they are indispensable to the learner himself, because we can regard the making of errors as a device the learner uses in order to learn.

(Corder 1967: 167)

Selinker (1972) also argued that the process of second language learning could be studied from three productive linguistic systems – the learner's native language utterances, his intermediary language competence in the target language (hence the term "interlanguage"), and the system of the target language. Accordingly, most studies, while trying to describe the systematicity of particular interlanguages, also tried to find what common paths of language development of L2 learners through analysing common error types and common sequences across interlanguages. In the 1970s, many researchers (Brown 1973; de Villiers and de Villiers 1973; Dulay and Burt 1973, 1974a, 1974b; Bailey, Madden and Krashen 1974; Larsen-Freeman 1975; Kessler and Idar 1979) carried out studies to examine learners' creative processes regarding the development of grammatical systems, hence often known as creative constructions studies. These studies

are predominantly based on the examination of grammatical morphemes aiming to describe the natural order of acquisition, therefore also referred to as morpheme studies.

For example, based on the findings from their two research studies on child second language acquisition involving error analysis and comparison of the acquisition sequence of certain grammatical morphemes, Dulay and Burt (1973) suggested that children's acquisition of second language syntax could be principally explained by their innate ability to organise structure in natural communication situations. They found (1974a) that Children's errors made when learning a second language in these situations were mainly accounted for in terms of "developmental strategies" (87.1%) rather than "first language interference" (4.7%). Dulay and Burt (1974b) confirmed the universal developmental learning strategies in child second language acquisition in another study, in which the acquisition sequences of eleven English functors were compared and analysed between children of Chinese- and Spanish-speaking background learning English.

However, like the Contrastive Analysis, the Error Analysis is also target-oriented (i.e. either the native language or the interlanguage are compared against the standards in the target language) and unable to adequately address what learners are doing correctly and appropriately and how learners avoid some target constructions so as to produce fewer errors. The methodology employed by morpheme studies (e.g. suppliance in obligatory context analysis) is criticised as striving for systematicity while obscuring variability in individual performances (Rosansky 1976; Andersen 1978). Moreover, these studies, on the whole, overlook contexts and functions of the linguistic forms, thus leaving learners' learning process unaddressed, a stance which variationists take in their research.

2.2.1.3. Monitor Theory

Based on the results of second language research by himself and other researchers, Krashen (1977, 1982) put forward his "Monitor theory" which assumed a common underlying route of acquisition for both L1 and L2 development. In his Monitor model, Krashen assumed learning and acquisition as two distinctive processes in language acquisition with the former leading to a consciously 'learned system' – explicit

knowledge, and the latter to a subconsciously 'acquired system' – implicit knowledge. He argued that the acquired system could act as a generator of spontaneous utterances in a situation of communicative demand, while the learned system could work only as a check or monitor on the output of the acquired system under certain very limited conditions and could not be converted into acquired knowledge.

For example, Krashen (1982) suggested that grammar has only two limited roles in second language instruction either as a monitor which "is limited to easily learned, late-acquired rules, simple morphological additions that do not make an overwhelming contribution to communicating the speaker or writer's message" (1982: 112), or as subject-matter through which the students' possible progress in SLA "is coming from the medium and not the message" (1982: 120). Therefore, "Neither role is essential, neither is the central part of the pedagogical program" (1982: 89).

In order to reduce the undesirable effect of monitor and to promote the process of language acquisition, (Krashen 1977) proposed the Input Hypothesis in which crucial to the acquired system was the comprehensible input, the input of language that is a little beyond the current level of the learner's comprehension. Krashen argued that the formal structure of language does not need to be explicitly learned or taught and will be automatically provided once there is sufficient comprehensible input and the learner understands it and communicates successfully with others. On the basis of these basic assumptions Krashen and Terrell (1983) put forth *the Natural Approach* to address second language instruction, the underlying assumption of which is that language acquisition occurs by understanding messages rather than formal structures, and that the major objective for language instruction should be to provide comprehensible input.

Though advocating a different theory from his contemporaries, Krashen and his Monitor theory "owed much to early interlanguage theory and to the research that it spawned" (Ellis 1994: 355). Krashen's theory, like the Error Analysis and the Interlanguage Theory, is also influenced by the rationalist view of language in that the mentalist notions such as innate linguistic knowledge and universal grammar are inherent in his

learning/acquisition, monitor, natural order hypotheses, albeit not quite so in his input hypothesis. White (1987) regarded Krashen's input hypothesis as inadequate due to its focus on meaning and context regardless of some context/meaning-independent aspects of grammar development, and due to its overemphasis on the role and benefits of simplified input without precision to identify what aspects of input are relevant to what stage of development. White (1987) argued for the inclusion of such an input hypothesis in L2 acquisition theory on the basis of having these problems adequately addressed. However, Krashen's theory has attracted some serious criticism (e.g. Gregg 1984) for lack of rigour as well as his lack-of-evidence theory-hypothesising manner.

2.2.1.4. Variation (and Functionalism)

Pidgin-creole studies and sociolinguistic studies have influenced interlanguage research in the direction of variation. Although engaged in research in different areas, early prominent researchers such as Bailey (1971), Bickerton (1971) and Labov (1969) all argued for a valid theory of language on the proposition that language is inherently variable and this variability is systematic. In a variationist perspective, interlanguage studies should reflect respects of variation and change through time and space in learners' interlanguage. Therefore, SLA research conducted within a variationist paradigm is mainly concerned with how learners' interlanguage varies and changes as a function of different contexts such as social, cultural, situational or task contexts in which a second language is being learned/acquired.

Based on the findings of pidgin-creole studies as well as his own research, Schumann (1976, 1978, 1986) developed the *Pidginisation/Acculturation Model* in an attempt to address the interlanguage development of second language learners. Schumann (1978) argued that the early stages of SLA reflect 'pidginisation' in that both are characteristic of the simplification of language for being restricted to strictly communicative functions, and that the later stages of SLA reflect 'decreolisation' in that both are characteristic of the complication of language for being able to perform integrative and expressive functions. Schumann (1978) then elaborated the pidginisation model in arguing that social and psychological distance from the target language group causes persistence of

pidginisation in the speech of an L2 learner. That is, the greater the social and psychological distance there is between the learner and the target language group, the less successful the learner is in acquiring the target language, hence the more persistent his pidginised target language is. This was exemplified in a longitudinal case study, in which Schumann (1976) found that the lack of development in English acquisition of a 30-year-old learner over ten-month period was the result of pidginisation, i.e. the result of the learner's social and psychological distance from native speakers of English. Schumann (1986) clustered the social and psychological variables affecting SLA into a single variable, acculturation, hence the Acculturation Model, which predicts how successful learners acquire the target language is dependent upon the extent to which they acculturate to the target language group. Since the 1990s, Schumann's focus of research has shifted from incorporation of a cognitive component in his pidginisation/acculturation model (e.g. Schumann 1990) to a more recent concern about psychological and neurological basis of language learning (e.g. Schumann 2001).

The pidginisation/acculturation theory seems to imply that social factors are the direct cause of success or failure in learning an L2, which seems to offer only part of the picture of language acquisition in a social setting. It cannot adequately explain why some learners do succeed in learning an L2 in an adverse situation in terms of social distance, and why other learners have their L2 remain pidginised even when they have stayed for years in a setting where there is no social distance between these learners and the target language group. The success or failure of the learner in acquiring an L2 depends on the interplay of a number of factors such as motivation and maturation apart from Schumann's social factors.

Based on the examination of the data from several studies, Tarone (1983) showed that interlanguage speech production varied systematically with elicitation tasks and argued that the continuum paradigm could better account for the phenomenon of variability in interlanguage systems than a Chomskyan paradigm and the monitor theory. The systematic variability related to tasks was also shown in another study (Tarone 1985) in which the performance of L2 learners on a certain task (e.g. a written grammar test)

varied from their performance on another (an oral interview). Moreover, the task-related variability also correlated with the variability of styles used in the tasks, that is, when the tasks were ordered in terms of degree of attention to language form required, the styles used by L2 learners in response to those tasks could also be ordered along a continuum. This continuum paradigm in addressing systematic variability in interlanguage systems was more thoroughly developed in Tarone (1988), where she postulated a 'capability continuum' of styles that range from the 'vernacular' (which demands the least attention to language form) to the 'careful' style (which demands the most attention to language form). In L2 production, learners will vary in using different styles required by different degrees of attention to linguistic forms under different conditions of language use. When new forms enter any one of the styles along the continuum, they may also permeate into the other styles. Tarone (1988) suggested that styles towards the careful end of the continuum are more target-like and less systematic than those towards the vernacular end, which are in turn more internally stable and less permeable to invasion from other styles. Consequently, the learner's language capability develops as entry and spreading (usually towards the direction of careful style) of new forms occur.

Ellis (1985) interpreted interlanguage variability as having two dimensions – horizontal and vertical – the former referring to the learner's L2 knowledge at the specific point in time and the latter to the stages through which interlanguage evolves. Ellis (1985, 1992) argued that interlanguage is at once systematic and variable, and that systematic horizontal variability is a precursor of vertical growth in interlanguage. He views the systematic variability of learners' interlanguage as a positive phenomenon, because it is the essential evidence of growth of an inherently unstable system still being developed. To address such systematic variability, unlike Tarone, Ellis (1985, 1992) suggested that learners have just one 'style' or grammar, but rules within this grammar may be variable. In the development of the learner's interlanguage, rules are initially constructed for unique use. Gradually, different forms associated with the same rule may emerge to compete for entry into the interlanguage system resulting in free variation, and finally each variant of the same rule may be restricted to a specific context of use or a specific grammatical function. Thus a form-function network in which different forms are

employed to perform different functions is built, and the development of the learner's L2 system occurs.

According to Towell and Hawkins (1994: 42), Ellis's and Tarone's theories of variation, like Schumann's pidginisation/acculturation hypothesis, suffer "from the absence of a theory of the nature of grammatical structure". Moreover, in view of the variation theories of Ellis and Tarone, Towell and Hawkins (1994: 43) made the following critical comments:

The problems that were found with these accounts are these. Transfer from the L1 receives no explanation. New forms enter the grammar freely (randomly), but this is in conflict with the considerable evidence for systematic staged development, which would be inexplicable if new forms freely entered the grammar. No account is offered for why some kinds of variation are resolved more quickly than others.

Under the influence of sociolinguistics, there has been a tendency for SLA research to be conducted in naturally-occurring language environments in order to examine how interlanguages are developed in natural settings and what impact social context has on acquisition such as in longitudinal case studies (e.g. Huebner 1983; Kumpf 1983; Sato 1990). Research along this line is based on naturally-occurring data, and if data elicitation does not occur in natural settings, attempts are made to replicate natural conditions to the extent possible so as to achieve generalisations in more contextualised terms.

Recently, Chinese learners' interlanguage variation has been examined in a number of studies (e.g. Young 1991; Bayley 1996; Wei 1997). Based on a systematic study of variation in interlanguage morphology of Chinese learners, Young (1991) proposed a new descriptive model for handling what the previous investigators had claimed to be random variations in interlanguage performance. Using the technique of VARBRUL to examine the data regarding plural inflections in the English interlanguage of Chinese learners, Young (1991) found that a multi-factored explanation rather than a single-

factored explanation may account for interlanguage development more effectively. In terms of statistical significance, his model tries to account for how factors such as the social context of speech, the linguistic environment of a variable, and the tendency to omit redundant information affect the developing interlanguage system.

Bayley (1996) used the technique of variable rule analysis to examine both linguistic and social factors that constrain surface realisation of underlying forms in the interlanguage of Chinese learners of English and thus to estimate the degree to which their production mirrors their acquisition. This technique provides a multidimensional picture of the English interlanguage of Chinese learners in terms of *-t/d* deletion and affixation: their interlanguage shows convergence with native varieties of English in phonological environment and speech style and divergence from native varieties in syllable stress, cluster length and grammatical category. Their English proficiency and social network affiliation are also shown to play a role in their production. The interaction of competing constraints of this technique is adequately addressed to account for the multiple factors that affect second language production and the technique is thus claimed to provide an effective means of measuring acquisition of *-t/d* deletion and affixation in the interlanguage of Chinese learners.

2.2.1.5. Universal Approaches

Since the 1980s, considerable attention has been attached to explaining facts about SLA in terms of universal properties of language. According to Gass (1989), there are three different approaches towards language universals: one advocated by Chomsky (1981a), one by Greenberg (1963, 1976) and the other by Bates and MacWhinney (1982). The approach advocated by Chomsky and his fellow researchers known as the Universal Grammar, argues that language universals, which reflect the innate properties of the human mind, underlie any human language and can be deductively discovered by studying a limited number of individual languages and consistently explained within the formalist domain in terms of abstract formal constructs of syntactic structures.

The typological approach discussed above (see 1.2 for details) claims that language universals should be formulated on the basis of observations across a wide range of languages, and that thus inductively-formulated language universals can be explained from language-externally motivated consideration (such as discourse and processing factors) as well as language-internal account (such as structural properties). The third approach developed by Bates and MacWhinney deals with language universals from a psychological perspective, or more specifically, a processing perspective. Their Competition Model (Bates and MacWhinney 1982) is initially developed to address how monolingual speakers interpret sentences and some attempts have since been made to investigate to what extent this model can be applied in accounting for SLA issues.

Since the typological approach of language universals and its implication for and application in SLA have already been illustrated in sections 1.2, 1.3, 1.4 and 2.1, only the other two universal approaches are briefly reviewed below.

Universal Grammar of the Chomskyan school (e.g. Chomsky 1981a) consists of a set of abstract and general principles, which is assumed to underlie the core grammars of all natural languages. The set of principles is postulated to represent the human beings' innate language faculty with which children acquire their native languages relatively rapidly and uniformly. The observed systematic variability found among languages can be accounted for by a set of well-defined parameters constrained by the principles. The invariant principles form the basis for all natural language grammars, but the parametric variation of the principles account for both the specified limits of human languages and the relative rapidity and uniformity with which all children are entitled to acquire all natural languages.

This principles-and-parameters approach of Universal Grammar has inspired a significant amount of SLA research, which differs to varying extent with respect to the accessibility of UG (evident in L1 acquisition) to L2 acquisition. Some (e.g. Epstein, et al. 1996, 1998; Flynn 1984, 1987, 1996; White 1985; White and Juffs 1998) have argued for a *full access* hypothesis which claims that the principles and parameters of Universal Grammar are

available not only to the child L1 learner but also to the adult L2 learner. This approach claims that differences between patterns of L2 acquisition and L1 acquisition can be explained in terms of other factors rather than a lack of access to Universal Grammar. Others (e.g. Bley-Vroman 1989; Clahsen 1988) argue for the *no access* hypothesis which claims that the principles and parameters of Universal Grammar is completely inaccessible to the adult L2 learner, whose acquisition of a second language can be explained in nonlinguistic terms such as types of problem-solving strategies. Still others (e.g. Schachter 1989) argue for the *partial access* hypothesis, claiming that the adult L2 learner has partial access to the principles and parameters of Universal Grammar only if they are instantiated in his/her L1 grammar. They also claim that other than that L2 acquisition should be explained in terms of problem-solving strategies as well.

For example, Flynn (1984) argues that the primary branching direction, which provides a significant constraint on first language acquisition in early stages, also holds in SLA. Based on her experimental data from two different L1 groups: Spanish (whose primary branching direction is right-branching - the same as English) and Japanese (whose primary branching direction is left-branching), she suggests that her results, which cannot be explained adequately in terms of either Lado's (1983) Contrastive Analysis (which assumes that L1 experience is important in SLA) and Dulay and Burt's (1974a, 1974b) strict Creative Construction Theory (which assumes that principles of acquisition independent of L1 experience are essential to SLA) can be best accounted for in terms of parameterisation of Universal Grammar principles. The primary branching direction is proposed as a significant principle in SLA: SLA is facilitated when parametric values of primary branching direction between the first and second language do match; while it is delayed when there is a mismatch in the parametric values. In another study using again two groups of English learners of Spanish- and Japanese-speaking backgrounds, Flynn (1987) investigates the role of the head-initial/head-final parameter in their acquisition of English pronoun anaphora. Given the results that both groups of learners use the head-initial/head-final parameter as a source of structural organisation for the L2, Flynn (1987) argues that the parameter-setting model of Universal Grammar for SLA is better positioned for accounting L2 acquisition phenomena in that what both the Contrastive

Analysis and the Creative Construction Theory try to explain is well addressed in this model.

The full-access-to-Universal-Grammar approach is more thoroughly expounded in Flynn (1996), and is further evidenced in Epstein, Flynn and Martohardjono's (1996, 1998) recent studies in which different levels of Japanese speakers learning English were examined via elicited imitation tasks and results show that all functional categories of Universal Grammar are available at every stage of Japanese learners' L2 acquisition and the observed errors are attributed to problems in production, not to lack of knowledge of a functional category, thus lending support to the strong continuity hypothesis. The corroborative evidence for the accessibility of functional categories of Universal Grammar also comes from White and Juffs (1998) and Prevost and White (2000). The former is a comparative study on native Chinese (a *wh*-in-situ language) speakers' and native English speakers' (as controls) acquisition of Universal Grammar constraints on *wh*-movement in English in a timed grammaticality judgment task and an untimed written question formation task (both designed to test long-distance *wh*-extraction under different island conditions). Based on the results that in both tasks Chinese learners of English are highly sensitive to subjacency violations (very few *wh*-in-situ questions: 5%-7%) and even no less so than English-speaking controls (27%), White and Juffs (1998) argue that even when they are exposed to the target language only in their native country, L2 learners still have access to universal grammar. In the latter study, the full-access-to-Universal-Grammar-based missing surface inflection hypothesis was examined against the impaired representation hypothesis in the spontaneous production data from two adult learners of French and two adult learners of German.¹⁷ The data show that although learners sometimes have a problem with realisation of surface morphology, such that they resort to nonfinite forms, generally finite forms do not occur in nonfinite contexts, learners exhibit syntactic reflexes of finiteness, and inflected forms largely show accurate

¹⁷ The missing surface inflection hypothesis means that L2 learners have unconscious knowledge of the functional projections and features underlying tense and agreement, while the impaired representation hypothesis means that L2 inflection is essentially impaired, due to lack of functional categories, features, or feature strength (Prevost and White 2000).

agreement. Prevost and White (2000) suggest that adult L2 learners represent finiteness and agreement at an abstract level, rather than being impaired in this domain.

There is a notable commonality between the generative grammar and typological approach in their respective application of markedness theory for predicting language acquisition. Croft (1990: 64) claims that although both the generative and typological approaches have adapted the concept of markedness from the Prague School of linguistic theory, "markedness in generative grammar is considerably different from markedness in typology". However, White (1989) argues that as regards some of their respective markedness claims, the two approaches are compatible in some of the areas they both identify as important, and in their similarities at the level of acquisition prediction for either first or second language. Mazurkewich (1984) and Zobl (1995) following generative approach, and Eckman (1977, 1996) following functional-typological approach, both make reference to the notion of markedness in their respective account for SLA phenomena with the similar prediction – the unmarked forms are usually acquired earlier or more easily than the marked forms. For example, Mazurkewich (1984) used two groups of English learners: speakers of Inuktitut and French for her study of the issue of markedness in the acquisition of dative questions. She hypothesises that the unmarked form of dative questions, the one with the preposition pie-piped with *wh*-word would be acquired prior to the marked form, the one with the preposition stranded in situ, despite the fact that the unmarked form may occur quite infrequently in the learner's language input. An operational written test in which subjects are asked to question the underlined phrases in a series of declarative sentences was administered and results confirm her prediction that the unmarked dative question is acquired before the marked one. So we might conclude that the view of full-access-Universal-Grammar proponents, is consistent with the view of typological universals that "All universals that are true for primary languages are also true for ILs" (Eckman 1996: 204).

In their model, Bates and MacWhinney (1982) presented two levels of language: form and function, that is, the forms of language are used to express communicative intentions. They argue that speakers use cues in determining relationships among elements. Central

to the argument is that while the range of cues is universal, the realisation of cues and the strength assigned to specific cues are language-specific. When there are elements of a sentence determined by cues which do not converge to the interpretation of the sentence, competition occurs as to which element will fill which grammatical position. Importantly, different languages resort to different cues in resolving the conflict. For instance, Gass (1989: 519) argues that "English uses word order and agreement as primary determinants. Other languages, for example, Italian in which word order is more flexible, rely more heavily on morphological agreement as well as on semantics and pragmatics." Based on the results of previous studies conducted cross-linguistically to investigate the interacting cues of word order and animacy, Gass (1989) concludes that in languages that depend on animacy for a basic interpretation strategy, speakers of those languages do not resort to a strong word order cue.

In a more recent longitudinal study of three German-English bilingual children in Australia, Döpke (1998) examined their developmental paths in acquiring word order in the verb phrases in German and English. She analyses her data in the light of cue strength and cue cost associated with each language. She finds that the major differences between these bilingual children and monolingual German- and English-speaking children are those concerning word order and finiteness of verb phrases in their German prominently in Phase II, and these differences can be best explained by the theory of cue competition. Döpke claims that her findings "provide support for cue competition as interface between input data and UG operations" (1998: 581) and deepen our understanding of the process of language acquisition from a new perspective, which cannot be otherwise achieved by studying monolingual language speakers alone.

Apart from the studies discussed above, there have been numerous other attempts at addressing SLA phenomena from the universal perspective either influenced by one of the above three approaches identified above or developed in their own right. Seliger (1984) proposed a model of processing in second language acquisition at two levels: strategy and tactic. The former refers to an innate biological mechanism which enables human beings to process information received (whether from L1 or L2 and whether from

the limited or elaborate system) in the same way, that is, through hypothesis formation and testing – the mind contrasting the known with the unknown and attempting to test, confirm or reject the unknown by relating it to the known. The latter involves individual learners, their learning situations and their linguistic responses in these situations. The strategy is universal and the tactic is specific. Even though it is expected that all individual learners resort to the same universal strategies, individual differences in L2 acquisition arise due to learners' choice of tactics. The selection of preferred tactics by individuals is, in turn, affected by the specific learning environment and learners' filter, which acts as the interface between strategy and tactic and determines what language data is selected for attention and learning (Seliger 1984: 47).

In reviewing the different ordering of certain English morphemes in both first and second language acquisition data obtained by Brown (1973), Dulay and Burt (1974c) and Bailey et al. (1974), VanPatten (1984) showed that the different orders obtained by these researchers were fundamentally the same if viewed in terms of separate syntactic positions (e.g., N-bound morphemes and V-bound morphemes). On the basis of this analysis and the notions of acquisition/learning distinction and monitor (e.g. Krashen 1980), VanPatten (1984) then went on to explain why these orders of morpheme acquisition for both L1 and L2 were the same. For example, within the acquisition order of V-bound morphemes for all learners of English (-ing > irregular past > -s),¹⁸ -ing is more communicatively important (for carrying more "semantic clout") than *irregular past*, which, in turn, is more communicatively important than -s, hence the order. In other words, -s is more likely to co-occur with other lexical items bearing the same information in the speech or is less importance for sentence comprehension than *irregular past*, which, in turn, is so than -ing. Therefore, the communication-based strategy works predominantly at early stages of acquisition and the learning strategy comes in only when learners become more proficient at meaning processing and begin to attend to less communicatively important forms (morphemes).

¹⁸ '>' means 'is acquired earlier than' and -s means 'third person singular' (see VanPatten 1984: 95-7 for a detailed illustration).

To sum up, research conducted within different paradigms in the field of SLA, varies in its theoretical orientations and empirical endeavours. Each approach contributes in its own right to the developing field. In view of the relationship between SLA and other disciplines, Gass (1993) argues that although SLA is already a field in its own right, SLA must not only draw from other disciplines, but must also provide intellectual stimulation to other disciplines so as to become a recognised field of academic inquiry with credibility.

2.2.2. Studies of Interlanguage in Question Formation: a Typological Approach

There have been a number of studies of English interlanguage relating to the acquisition of *yes/no* questions and/or *wh* questions by English learners of different L1 backgrounds in the literature of SLA (e.g. Mukattash 1981; Saunders 1983; Eckman et al 1989; Brines 1990; Kim 1999). These studies were carried out from different theoretical/pedagogical perspectives; different from others, Eckman et al (1989) was a study conducted in a typological approach.

In order to directly address the question of whether primary-language universals also hold for secondary languages such as interlanguages, Eckman, Moravcsik and Wirth (1989) conducted a study to examine the validity of the implicational universal pertaining to the word order of questions (Greenberg 1963) against interlanguage data. In the study, they used a task of eliciting questions around story squares from 14 English learners of Japanese- (N = 6), Korean- (N = 4) and Turkish-speaking (N = 4) backgrounds respectively. The gathered interlanguage speech data were then used for testing the implicational universal.

They reinterpreted the implicational universal pertaining to word order of questions – Greenberg's Universal 11 (see (3) in 2.1.3.1) – as the following two separate implicational universals (Eckman et al 1989: 175):

- (38) A. "*Wh* Inversion implies *Wh* Fronting"
B. "*Yes/No* Inversion implies *Wh* Inversion"

To operationalise the examination of the two universals, Eckman et al (1989: 180) first defined some basic grammatical constructions to which relevant utterances were tied, then tied these grammatical constructions to relevant general patterns, and finally related these general patterns to the implicational universals to be tested. After defining grammatical constructions such as *subject*, *verb*, *yes/no* and *wh inversion*, and classifying the actual instances accordingly, Eckman et al (1989: 175) established the following three patterns:

- (39) A. *Wh Fronting* – the sentence-initial (vs. non-initial) position of question words or phrases in *wh* questions ("What are you doing?" vs. "You are doing what?");
B. *Wh Inversion* – verb-before-subject (vs. subject-before-verb) order in *Wh* questions ("What are you doing?" vs. "What you are doing?");
C. *Yes/No Inversion* – verb-before-subject (vs. subject-before-verb) order in yes/no questions ("Are you going back?" vs. "You are going back?").

The relevant classified utterances were then related to one of the three patterns in (39), calculated and tabulated for analysis (see Eckman et al 1989: 181-5). Results of the subjects' performances, in compliance with the *Interlanguage Structural Conformity Hypothesis*,¹⁹ predominantly support the two universals with only one exception to 38B (see Eckman et al 1989: 185-7). It is significant that not only are the two implicational universals upheld in interlanguage data (Eckman et al 1989), but their findings also suggest that typological universals are indeed valid predictors for L2 acquisition.

This study, in a similar vein, will use a different task – the naturalistic conversation comprising an interview and a role-play – to further examine the predictive power of the two implicational universals for interlanguage data.

¹⁹ *Interlanguage Structural Conformity Hypothesis*: All universals that are true for primary languages are also true for interlanguages (Eckman et al 1989: 195).

2.2.3. Studies of Interlanguage in Relative Clause Formation

The acquisition of English relative clauses by L2 learners is a well-studied area and is well documented in the literature of SLA. This phenomenon has been studied by different researchers from different perspectives; and consequently, in addition to some common characteristics found in the relevant studies, the research findings may vary with respect to underlying theoretical assumptions and methodological issues involved in the research within the same research paradigm as well as across different research paradigms. What follows hereafter reviews briefly the studies of relative clause formation in SLA, and then in detail those studies relating to the typological universal – the Accessibility Hierarchy (Keenan and Comrie 1977).

2.2.3.1. Relative Clause Formation in SLA: an Overview

As has been demonstrated in the previous section (2.2.1), research in SLA or interlanguage has been strongly influenced by different theoretical paradigms in the disciplines of linguistics, psychology and sociology. There is no exception for studies regarding the acquisition of English relative clauses by L2 learners.

One of the early studies conducted in this area is Cook' (1973) comparison of language development between native children and foreign adults. In order to supply some concrete facts rather than the sole focus on theoretical speculation on this issue, Cook (1973), by adapting some procedures from developmental psycholinguistics, devised two experiments, one of which was a comparative study of the imitation and comprehension of relative clauses between these two groups. Based on the results that both groups seemed to share far more similarities than differences, Cook (1973: 20-22) concluded that adult L2 learning is fundamentally similar to child L1 learning, in contrast to 'the teacher's belief that foreign adults approach language in ways fundamentally different from native children'. However, a very important difference between foreign adults and native children (i.e. only children repeated the last few words of the sentence) was attributed to mental differences in terms of memory capacity instead of language differences. In effect, Cook (1977) confirmed this difference in another study, which

consisted of a series of experiments including a test of English relative clause comprehension among native English-speaking adults and children and foreign adults.

Tushyeh (1988) used an error analysis in examining Arabic-speaking learners' acquisition of English relative clauses in different tasks (translation test, grammaticality judgment, sentence combination and sentence completion). Results (Tushyeh 1988) showed that the use of the resumptive pronoun was the most frequently occurring error type, which was accounted for by language transfer, and that other error types included omission of the obligatory relative pronoun, preposition omission, and use of possessive pronouns with the antecedent. In addition, overgeneralisation of target language features and simplification were involved to address those intralingual errors.

Kupferberg and Olshtain (1996) investigated the effect of contrastive linguistic input on the learning of difficult English forms such as compound nouns and nonfinite restrictive relative clauses by native Hebrew speakers. In their study, natural linguistic input was provided for both experimental and control groups, but additional contrastive linguistic input was introduced only to the former. Based on the findings that the scores for recognition and production tasks for the experimental group learners were significantly increased with exposure to contrastive linguistic input, Kupferberg and Olshtain (1996) suggest that contrastive input expedites noticing difficult linguistic forms and should be incorporated into L2 programs.

Both the Universal Grammar-based principles-and-parameters approach and the typological universal approach have been widely applied in examining acquisition of English relative clauses by L2 learners. For example, Flynn (1989) investigated the role of the head-initial/head-final parameter in the acquisition of bound variables in English restrictive relative clauses by adult Spanish and Japanese speakers in an elicited imitation task. Based on her experimental results that significant differences in patterns of acquisition between Spanish and Japanese speakers did exist, she (1989: 102-5) argued that the differences can be explained in terms of the differences between the head direction of the L1 and L2 of the learners, and that the head-direction parameter

constrains SLA, which lends support to the parameter-setting model of L2 acquisition in particular and a theory of parameter-setting in the framework of Universal Grammar in general.

In another study based on the generative framework, Adjémian and Liceras (1984) examined the internal structures of restrictive relative clauses of English, French and Spanish and the problems these structures posed for L2 learners. The acquisition of restrictive relative clauses by adult native speakers of English learning French or Spanish and by adult native speakers of French learning English or Spanish was tested in two oral elicitation tasks and three written tasks. Results showed that the acquisition of some subtle differences in relative clauses between related languages could not be straightforwardly addressed with an unidimensional explanation; rather, the emerging interlanguage grammar of L2 learners was shaped by the interaction of several factors including transfer, universal grammar and learner-produced hypotheses (Adjémian and Liceras: 107-16).

Comparable to the vast amount of SLA research conducted in the generative paradigm is enormous amount of SLA research relating to the Accessibility Hierarchy in the typological paradigm, to which we now turn.

2.2.3.2. Relative Clause Formation in SLA: Accessibility Hierarchy

One of the questions addressed at the Conference on Language Universals and Second Language Acquisition held at the University of Southern California in February 1982, is: "What might the existence of accessibility hierarchies lead one to look for in the analysis of interlanguages?" (Rutherford 1984: 2) With respect to this question, Comrie has (1984: 14-5) observed that:

The cross-language study of relative clause formation presented in Keenan and Comrie (1977) has spawned a vast amount of relevant literature in the second language acquisition area, showing how the theoretical conclusions reached by Keenan and Comrie translate fairly directly into valid predictions about the

acquisition of relative clauses in a second language, though also noting more specific points where the fit between the two areas is less than perfect.

In effect, through examination of the relationship between language universals and language acquisition studies in relation to relative clause formation, Gass and Ard (1980) found that L2 acquisition data provide a clearer window for investigation and verification of language universals than do L1 acquisition data. L1 acquisition appears to be strongly affected by factors rooted in cognitive development, while adult L2 acquisition seems to be affected by factors primarily based on language universals. Indeed, since the inception of the Accessibility Hierarchy, the validity of the Accessibility Hierarchy in language acquisition has been challenged more frequently in studies of primary languages (e.g. Yeoh 1977; Manaster-Ramer 1979; Joseph 1983; Lewis 1985; Lawal 1987) than in L2 studies (e.g. Ito 2001). Most studies (e.g. Gass 1979, 1982; Pavesi 1986; Eckman et al 1988; Li and Li 1994; Sadighi 1994; Sadighi and Jafarpurin 1994; Aarts and Schils 1995; Park 1998, 2000) in SLA relating to the Accessibility Hierarchy, in general, seem to lend support to its validity in predictions about L2 acquisition, albeit not without controversies.

Gass (1979) was the first researcher to test the relevance of the Accessibility Hierarchy in SLA studies. In an attempt to better determine the nature of language transfer in relation to the role of universals of grammatical relations, Gass conducted a study which investigated the formation of relative clauses based on the Accessibility Hierarchy in SLA. The subjects were 17 adult L2 learners of English with 9 different native languages – Arabic, Chinese, French, Italian, Korean, Persian, Portuguese, Japanese and Thai. In order to test learners' receptive and productive linguistic knowledge of RC formation, the tasks of grammaticality judgment and sentence combination were used respectively. The former involved giving acceptability judgments to 29 English sentences with each containing a restrictive relative (13 are well-formed and 16 are not) and the latter involved combining two separate sentences into a single sentence containing a relative clause (with 12 pairs to complete).

There are a number of findings reported in this study. First, L1 transfer effects are only significant for the variable of *pronoun retention* but not in the other three: *relative clause marker omission*, *relative clause marker selection* and *adjacency*.²⁰ Second, target-language facts should be considered in the description of language transfer (Gass 1979: 342). Third and most important, the percentages of the L2 learners' sentences correct on the sentence combination task match neatly with the grammatical relations (with the exception of GEN)²¹ of the Accessibility Hierarchy, as shown in Table 2.2, hence validating this typological universal by the data in this study.

Table 2.2. Percentage of Sentences Correct by Language Groups

SU	DO	IO/OBL	GEN	OCOMP	Language Group
48	28	27	47	10	Arabic
90	30	25	30	13	Thai
68	28	17	33	0	Romance
70	20	18	38	4	Persian
78	60	28	88	0	Chinese, Japanese, Korean

(Adapted from Gass 1979: 340)

As Gass put it, language universals "were found to play the leading role in this study since they were dominant both in assigning relative orders of difficulty and in determining where language transfer occurs" (1979: 341). Based on the results of the study, Gass (1979) proposed a model of language transfer that predicts the conditions most conducive to transfer, including notions of language universals, language distance, and surface language phenomena.

In another study, Gass (1982) examined whether the Accessibility Hierarchy for SLA has some implications for L2 instruction, that is, whether instruction on relativisation on a low position on the hierarchy given to L2 learners can enable them to make

²⁰ For the grammaticality judgment task, the transfer effect is on the positions of SU, DO and IO; for the sentence combination task, it is on the positions of DO, IO and OBL (for details of explanation, see Gass 1979: 335-9).

²¹ Gass (1979: 341) provides two possible explanations for this. *Whose* in English is the only invariant relative marker coded in GEN position, hence being salient and easily perceivable by the L2 learner. Another possibility is specifically related to the study in which the learners may have interpreted the genitive marker and the noun following it as a whole unit, which in fact covers either SU or DO positions (positions high on the hierarchy) in all instances containing genitive markers.

generalisations to the higher but not lower positions on the hierarchy. She used two groups of L2 learners of English with Arabic, Italian, Persian, Russian and Spanish as their native languages respectively; one is the experimental group with thirteen learners and the other is the control group with five learners. Pre-instruction (grammaticality judgment and sentence combination) tests showed no significant difference between the two groups in terms of their pre-instructional knowledge of English relative clauses. Three days later, the experimental group received relativisation instruction on OBL position only, while the control group was given relativisation instruction from a textbook starting from SU, to DO, IO and finally to GEN (with less emphasised instruction) positions. About two days after the instruction, post-instruction tests on all the positions on the Accessibility Hierarchy were administered to both groups, and results showed that the experimental group had improved significantly in terms of score difference between pre- and post-instruction tests while the control group had not. Moreover, in the task of combining two separate sentences into one sentence containing a relative clause (production task), the experimental group demonstrated the ability to generalise from relativisation on OBL position to all other (SU, DO, ID, OCOMP) but GEN positions, whereas the improvement for the control group was mainly confined to what they had been taught. The results of the improvement of the two groups on the production task between the pre- and post-tests are presented in Table 2.3 in terms of percentages.

Table 2.3. Improvement on the production task in the two groups

Control group		Experimental group	
SU	40%	SU	30%
DO	30%	DO	39%
IO	0%	IO	42%
OBL	40%	OBL	57%
GEN	10%	GEN	12%
OCOMP	0%	OCOMP	50%

(Adapted from Gass 1982: 138)

A number of interesting points can be seen from the table. First, the control group did not improve at all for relativisation on IO position but did improve markedly for relativisation on OBL position for which they received no instruction. Second, the experimental group generalises from OBL relativisation to relativisation not only on

higher positions (e.g. SU, DO, IO) but on lower positions (e.g. OCOMP) as well. Third, both groups improve only marginally on GEN relativisation. This result runs counter to Gass's earlier findings (Gass 1979; cf. Note 19). Nevertheless, Gass's (1982) hypothesis is well supported by the overall results of the study, which lead her to draw a pedagogical implication for syllabus design in which she suggests difficult structures should precede the easier ones given learners' 'natural' generalising abilities (Gass 1982: 139).

Gass's (1982) hypothesis was further examined by Eckman, Bell and Nelson (1988) who replicated and extended the findings of the former in an attempt to test the generalisation of instruction in L2 learning. Based on their *Markedness Differential Hypothesis*,²² Eckman et al (1988) suggest "that it is the most marked aspects of a target language from which it should be possible for a learner to gain maximal generalization of his/her learning" (1988: 4). As regards the Accessibility Hierarchy (Keenan and Comrie 1977), the lower the position of relativisation, the more marked it is and if learned, the more other higher positions on the hierarchy it can generalise to in learning.

To test more finely for generalisation of L2 learning, Eckman et al (1988) used 36 ESL students who were divided into three experimental groups and one control group based on their native languages (four Arabic speakers, three Spanish speakers, one Japanese speaker and one Korean speaker for each group), their English proficiency level as well as their results of a pretest on combining two sentences into one sentence with a relative clause. Each experimental group was taught to form relative clauses on only one position on the Accessibility Hierarchy (i.e. SU, DO or OBL), while the control group was instructed not in relative clause formation but in sentence combining techniques unrelated to relative clauses. Two days after the instruction, all the four groups were given a post-test. The results of the post-test showed that the OBL group scored the best, with the DO

²² Eckman et al (1988: 4) defines this hypothesis as:

"The areas of difficulty that a learner will have with a given TL can be predicted on the basis of a systematic comparison of the NL and TL, such that:

- (a) those areas of the TL which are different from the NL and relatively more marked than in the NL will be difficult;
- (b) the degree of difficulty of any aspect of the TL which is different from the NL and relatively more marked than in the NL will correspond to the relative degree of markedness of that aspect;

group next followed by the SU group, and the control group last, and that the generalisation of learning generally went upward on the hierarchy – neither the SU group nor the DO group generalised to the OBL position, though the SU group generalised somewhat to the DO position. Despite the fact that Eckman et al (1988: 10-12) are unable to explain the learners' performance on SU and DO relativisation (no statistically significant difference) in terms of predictions made by the Markedness Differential Hypothesis and a larger impact of training on the OBL group than on the other groups, the overall results do seem to support the pedagogical hypothesis underlying this study as well as Gass (1982). That is, "maximal generalization of learning will result from the acquisition of relatively more marked structures" (Eckman et al 1988: 12).

Pavesi (1986) used the Accessibility Hierarchy to investigate English relative clause formation by L2 learners of English with Italian-speaking background. Forty-eight teenage learners (aged 14-18) were tested in a written task – sentence combination and an oral task – picture description, both required to use relative clauses. Results were analysed in terms of the relationship between markedness and intertask variability. Pavesi (1986) found that L2 learners' interlanguage development could be explained in a regular manner – proceeding from unmarked to marked. Markedness in fact accounted for both intertask variability (the written task showed more marked features than the oral task) and most production involving non-target-like rather than target-like performance.

Based on his previous research (Sadighi 1982) in which the acquisition of English restrictive relative clauses by Persian native speakers was investigated, Sadighi (1994) examined the similar phenomenon among 56 adult English learners of Chinese- (24), Japanese- (20) and Korean- (12) speaking backgrounds. Used in this study was a comprehension test based on Sadighi (1982) consisting 42 multiple-choice instances out of 28 relative clause types, which were formed in the manner of combining matrix clause head NPs (SU, DO, IO, OBL) with relativised NPs on the Accessibility Hierarchy (SU, DO, IO, OBL, GEN-SU, GEN-DO, OCOMP) respectively. Errors were then analysed to

(c) those aspects of the TL which are different from the NL, but which are not more marked than in the NL will not be difficult."

see whether the three universal factors – interruption, word-order rearrangement, and parallel function could account for the processing of relative clauses among these learners, whose native languages all have a relative clause order distinct from English (i.e. RelN vs. NRel).²³ Based on the results of the study and other findings in the literature, Sadighi (1994: 147) concludes that language universal factors predominantly constrain both L1 and L2 language acquisition even when L2 learners' native languages diversify in language specificities, and that linguistic universals are accessible to both children and adults, remaining operative maturationally.

Apart from the three universal factors in accounting for his data, Sadighi did not make any explicit reference to what extent his data support the Accessibility Hierarchy except for his implication for the difficulty associated with the GEN and OCOMP positions (1994: 146):

Another important finding is related to the sentences containing possessives and object of comparative particles. The breakdown of these sentences revealed that they were the source of a large number of errors. Out of 436 errors made 211 (48.4%) belong to these sentences.

However, this claim is somewhat problematic. A close look at his data (Sadighi 1994: 150) reveals that the 211 errors contributing to GEN and OCOMP positions came from the 56 learners' performance on 21 instances, while their performance on the other 21 instances (for SU, DO, IO and OBL in all) produced as many as 225 errors (51.6%). Regardless of the position of matrix NP heads, Sadighi's (1994: 150) data can be recounted and presented in Table 2.4.

²³ By *interruption*, "self-embedded sentences, that is, sentences with matrix clause NPs and VP separated by a relativizing clause are harder to interpret than a sentence in which the main clause is not separated by a relative clause"; by *word order rearrangement*, "the subordinate clause with an SVO word order is much easier to comprehend than the one with an OSV word order" in which the subordinate co-referential NP is an object; and by *parallel function*, "sentences with the coreferential NPs functioning as subjects or objects are easier to understand than sentences with non-coreferential NPs" (Sadighi 1994: 145).

Table 2.4. Number and percentage of errors on the relativised positions in Sadighi (1994)

Relativised NP Type	Total Number of Instances	Total Number of Errors	Average Error Percentage	Difficulty Ranking
SU	4	47	20.1	5
DO	6	48	14.3	3
IO	6	55	16.4	4
OBL	5	75	26.8	7 (hardest)
GEN-SU	6	44	13.1	1 (easiest)
GEN-DO	8	113	25.2	6
OCOMP	7	54	13.8	2

From the table, it is clear that only the GEN-DO position creates a large number of errors ranked at the difficulty level 6, while both GEN-SU and OCOMP positions produce a very small number of errors on average ranked at the difficulty levels 1 and 2 respectively. Overall, the results do not seem to support the Accessibility Hierarchy; nonetheless, they are well accounted for by the three universal factors. In effect, the overall ranking of SU at the level 5 is largely due to the number of errors in the SU-SU type (24; 42.8%) with the DO-SU type producing the least number of errors of all types,²⁴ a phenomenon that is well explained by the interruption hypothesis (see footnote 23). Finally, the three universal factors involved by Sadighi (1994) in addressing the issue under discussion can be effectively replaced by one single processing approach in terms of the principle of Early Immediate Constituents (Hawkins 1994) or filler-gap domain (Hawkins 1999) (see 2.1.3 for details).

In another study which examined the role of learner treatment in comprehending English restrictive relative clauses under formal instruction, Sadighi and Jafarpur (1994) find that although no significant relationship between the frequency of restrictive relative clause types in classroom texts and the difficulty order obtained from L2 learners' (104 Shiraz University students) on a comprehension test of relative clauses, the relative ease/difficulty does relate to the above-mentioned universal factors. Moreover, the difficulty order corresponds closely with the NP positions on the Accessibility Hierarchy (Keenan and Comrie 1977).

²⁴ In 'SU-SU' and 'DO-SU', the position before the hyphen refers to the NP head in the matrix clause, and that after it refers to the relativised NP position in the relative clause (see Sadighi 1994: 144).

The validity of the Accessibility Hierarchy in SLA is further evidenced in a couple of recent studies. Based on a comparative analysis of relativisation in Chinese and English and the Accessibility Hierarchy, Li and Li (1994) examined the degree of relativisation of five positions (SU, DO, IO, OBL, GEN) on the hierarchy in Chinese and English both as L1 and L2 among Chinese students of English (N = 48) and American students of Chinese (N = 9). Four tests were administered to them for composing five types of sentences with each containing a relative clause and representing each of the five positions on the hierarchy. The overall results of the study support the Accessibility Hierarchy quite well; the results for the American students of Chinese (tests 3 and 4), in particular, match neatly with the hierarchy. Interestingly, the Chinese learners of English performed similarly to the L2 learners in Gass (1979) in GEN relativisation: the former treated GEN as more accessible than IO and OBL positions but less so than SU and DO positions (Li and Li 1994: 68-71), while the latter treated GEN as more accessible than DO, IO, OBL, and, occasionally, than SU positions by Chinese, Japanese and Korean speakers (Gass 1979: 340). However, Li and Li's (1994: 73) account for this is less convincing than the explanation supposed by Gass (1979: 341; see footnote 21).²⁵

Aarts and Schils (1995) conducted a study at the University of Nijmegen in order to examine the three factors - effectiveness of lectures, the Accessibility Hierarchy, the contrastive analysis hypothesis - in relative clause formation. Using sentence-combining tasks, they tested 96 first-year Dutch students of English twice with a 3-month interval, during which they had three 45-minute lectures on relative clauses. Results did not lend support to the contrastive analysis hypothesis regarding interference from the first language (Aarts and Schils 1995: 55), but the students' performance after the instruction had been greatly improved - 75 students (78.1%) scored better in the second test than in the first (p. 50). Moreover, their test results support the Accessibility Hierarchy (except for the order of SU and DO) and the claim made by Comrie (1981: 161) that the same

²⁵ The inaccuracy of Li and Li's account for their subjects' performance on GEN position can be seen in their statement "This may be because to the Chinese the English GEN is simpler than IO and OBL. IO and OBL in English have a complicated system of relative pronouns (which/that/who(m)), but GEN has just *whose* which is very close to the Chinese system of relativization in terms of the simplicity of relative pronouns." (1994: 73) In fact, DO in English also has that same complicated system of relative pronouns,

constituents are more difficult to relativise in a subordinate clause than in a main clause (Aarts and Schils 1995: 52-4). The absolute hypothesis of Fox (1987) may help to explain the preference of DO over SU on the hierarchy in this study and lack of differentiation of the two positions in the results of Eckman et al (1988: 11).

In two recent studies, Park (1998, 2000) again confirms the validity of the Accessibility Hierarchy in SLA studies concerning Korean learners of English. In the first study (Park 1988), the results of the tasks of sentence combination and grammaticality judgment show that L2 learners' interlanguage is constrained by universal grammar rather than their native language and that the order of difficulty in relativisation corresponds to the accessibility hierarchy. In the second study (Park 2000), two experiments using sentence combination, were administered to Korean learners of English at different levels (middle school: 33 & 37; high school: 33 & 37; college: 25 & 27 respectively for the two tests), and the findings confirm the Accessibility Hierarchy. In addition, the assertion that universal learning principles significantly influence interlanguage development is also well supported by the study.

As has been discussed in this section, the Accessibility Hierarchy, as a typological universal, has lent much to SLA studies, and will do alike in the future. In his closing remarks on the relation between language universals such as the Accessibility Hierarchy and second language acquisition, Comrie (1984: 27) stated:

Language universals of this kind lend themselves to ready transposition into predictions about ease of acquisition, and the large number of such universals that have been proposed in recent theoretical work should provide a fruitful field for interaction between language universals research and second language acquisition research.

In a similar vein, a number of recently-formulated universals by Hawkins (1999) as well as some existing typological universals in the literature such as the Accessibility

but is more accessible than GEN in their study. Nonetheless this conflict can be resolved by Gass's (1979: 34; also note 15) two plausible proposed explanations.

Hierarchy (Keenan and Comrie 1977) and question order prediction (Greenberg 1963), were tested in the interlanguage data of the present study, the methodology, results, major theoretical findings and implications of which are discussed in Chapters 3, 4, 5 and 6 respectively.

Chapter Three Methodology

This chapter consists of three sections. Section one (3.1) – the study – introduces the background of the study, procedure of data collection and participants, and difficulties encountered. Section two (3.2) – the tasks – gives a detailed description of the sampling tasks including a conversation (an interview and a role-play), repetition, sentence combination, grammaticality judgment and free essay writing. Section three (3.3) – data processing – involves the procedure for transcribing, coding, tabulating and sorting out data for data analysis.

3.1. The Study

3.1.1. Background of the Study

The present study is informed by the theoretical paradigm of typological universals.²⁶ This project examines a number of typological universals in interlanguage data. The universals examined include word order of interrogative questions (Greenberg 1963), which was examined in L2 data by Eckman et al (1989); the Accessibility Hierarchy (Keenan and Comrie 1977), which has been the subject of a number of different SLA studies; and some recently formulated universals (Hawkins 1999), which have not been researched from the perspective of L2 acquisition. Seliger and Shohamy (1989: 24-41) proposed four parameters essential to L2 research – a) synthetic and analytic approaches and b) heuristic and deductive objectives at the conceptual level, c) degree of control and manipulation of the research context and d) degree of explicitness in data collection procedures at the operational level. In terms of these parameters, the design of this study is characteristic of an analytic approach, deductive objectives, a fairly high degree of control and manipulation of the research context and a fairly high degree of explicitness in data collection procedures.

By the ‘analytic approach’, this study focused on investigating the validity of typological universals such as question formation (Greenberg 1963), the Accessibility Hierarchy

(Keenan and Comrie 1977) and some other implicational universals (Hawkins 1999) in SLA analytically and independent of each other though at some stage the 'holistic approach' was adopted as well. By the 'deductive objective', this study aimed to test such specific typological universals with respect to their explanatory value for understanding corresponding L2 phenomena though the 'heuristic purpose' was involved in observing and finding out systematic interlanguage features in the naturalistic data of the study (i.e. the conversation). By having a 'fairly high degree of control and manipulation of the research context', this study restricted its research focus to testing specific universals, exercised control over the variables involved, focused more on form in the tasks of elicited repetition, sentence combination and grammaticality judgment than in the conversation task (more on communication), and consequently involved a low degree of researcher subjectivity except for dealing with the conversation task (see Appendix III for the tasks used in this study). Finally, by having a 'fairly high degree of explicitness in data collection procedures', the data collection procedures of the study drew the subjects' attention to the procedures themselves more explicitly in the tasks of elicited repetition, sentence combination and grammaticality judgment than in the conversation task. The theoretical framework – a typological universal approach and the four parameters characteristic of the study – lay the basis for the whole research project.

3.1.2. Procedure of Data Collection and Participants

The design of the study, involved the collection of both comprehension and production data from adult ESL students. The comprehension data were collected via tasks of elicited repetition and grammaticality judgment, while the production data were collected via a sentence combination task as well as a near-naturally-occurring conversation and free essays. These tasks are described in detail in view of their design, purpose, the kind of data sought, perceived desirable and undesirable effects associated with the design and issues of improving their reliability and validity (see 3.2 for details). The data collection was carried out twice within a six-month span in an attempt to capture the impact of certain developmental features on the shaping of L2 learners' interlanguage in general

²⁶ I owe much to my supervisor, Dr Heather Bowe, for her guidance with expertise in my adopting the typological framework for the present research.

and the typological universals tested in the study in particular.²⁷ The data collected were then transcribed, coded, tabulated and sorted out for analysis (see 3.3 for details), and analysed both quantitatively and qualitatively (see Chapter 4 for details).

The ethical appropriateness of the study was checked through the Standing Committee on Ethics in Research Involving Humans (SCERH) of Monash University. The application for ethical clearance of the research was submitted to the committee and the final approval was obtained on September 11th 2000 (Project 2000/351).

A native English speaker was recruited to administer the interview and the task of cued elicitation of questions. The native English speaker recruited was middle-aged, agreeable, eloquent and, above all, very experienced and responsible. She had some experience of teaching English to foreign students, working as an attaché abroad and had been involved in some other communication-oriented work. It did not take long for her to get well acquainted with the interviewing job after training. Furthermore, a pilot study was conducted to try out the procedure of data collection before the commencement of the real study. In the pilot study, both the researcher and the native speaker were familiarised with administering the procedure and using the research instruments, and the problems that arose from the pilot study were resolved in the revised procedure of data collection.

The subjects recruited were international students studying various courses at the Monash University English Language Centre. The centre offers training programs in both General English (GE) and English for Academic Purposes (EAP). The former consists of six levels of courses (beginner, preliminary, pre-intermediate, intermediate, upper-intermediate and advanced, i.e. GE 1 – GE 6). The latter, commencing at least at the upper-intermediate level, comprises three streams – ETP (English for tertiary/test preparation), ELBP (English language bridging program) and IAP (Introductory academic program), all of which require different entry IELTS scores. While ETP levels

²⁷ Unfortunately, only five of the subjects recruited in the first round of data collection, participated in the second round. By the time of the commencement of the second-round data collection, many of the first round participants had already been enrolled in various degree courses at Monash University; some (e.g.

are compatible to GE higher levels (GE 5 & 6). ELBP and IAP levels are even higher representing respectively the near pass and pass of university entry requirement score for English proficiency.²⁸ The students enrolled in ELBP or IAP programs were excluded from the sampling population on two accounts. First, their levels of English proficiency were incompatible with the levels of those enrolled in GE programs. Furthermore, they attended types of university course-oriented intensive training for only five to fifteen weeks and did not need to sit an IELTS test (the speaking component of which the designed conversation simulated), hence they were excluded from recruitment. As a result, only GE and EPT students were targeted for sampling. These students were normally placed into classes of different levels according to their scores of the diagnostic placement test (which included an interview as well as listening, reading, grammar and writing sections) conducted upon their enrolment. Each class was usually made up of 14 to 16 students. Changes for class levels usually occurred at the beginning of each course and at the beginning of week six.

The students were recruited on a volunteer basis without being paid. In the targeted classes copies of the explanatory statement and the informed consent forms for the research were distributed to the students and explained to them in detail. If they agreed to participate, they signed and dated the informed consent forms and gave them back to the researcher. Then the subjects chose a specific time slot in the timetable for participation, in a 35-minute data collection process which included a conversation (an interview and a task of cued elicitation of questions) with a native English speaker and the tasks of elicited repetition, sentence combination and grammaticality judgment. The conversation

some Japanese subjects) had finished their course and left Australia; a few others were not interested in being enlisted again in the second round.

²⁸ International students whose native language is not English have to satisfy English language requirements so as to be accepted with a full offer in degree programs at Monash University. The minimum test score requirement set for undergraduate courses is: TOEFL 550 with a TWE (Test of Written English) score of 5 (Computer-based TOEFL: 213 with an Essay Rating (ER) score of 5, or IELTS-academic 6.0 with no individual band score less than 6.0. For all the postgraduate courses and some of the undergraduate courses offered in the faculties of Business and Economics, Medicine, Nursing and Health Sciences, Law and Pharmacy, the requirement is higher: TOEFL 577 with a TWE of 5 (computer based TOEFL: 233 with an ER of 5), or IELTS-academic 6.5 with no individual band score less than 6.0. However, students who do not meet the entry score in these tests can be accepted with a conditional offer in the first instance and given a full offer later provided that they have satisfied the English language requirement. They can

and the repetition tasks were audiotaped and the sentence combination and grammaticality judgment tasks were done on paper. After that, each subject was given a stamped return envelope with twenty cents enclosed (for copying cost) and asked to mail to the researcher a copy of an original essay they had done in class. The whole data collection was carried out in pre-arranged venues (classrooms) during after-class hours from 3:45 p.m. to 5:30 p.m.

In order to recruit sufficient representative subjects, data collection was carried out in two rounds with a six-month interval. It was also hoped that some first round participants could be recruited in the second round and that their interlanguage developmental features could be sufficiently examined after the proposed time span. In the first round, 25 out of the 30 recruited student volunteers participated in the data collection, and in the second round, 38 out of the 52 took part. As a result, the total number of participants were 63,²⁹ out of whom 60 participated in the conversation, 60 in the task of elicited imitation, 56 in the task of sentence combination, 56 in the task of grammaticality judgment, and 45 free essays were collected.

The students recruited at the centre fell broadly into three groups. A) Many were accepted in different degree programs at Monash University but with a conditional offer. B) Some did not have a university offer but wanted to improve their English through training either for future university study or for some other purposes (e.g. work or migration). There was also a special group of Japanese university students who studied English for six months as part of their course requirement.

Of the actual 58 subjects, 27 were male and 31 were female and their age ranged from 18 to 41 averaged as 23.8 years. They came from different countries speaking a variety of native languages: Mandarin Chinese (N = 15), Japanese (N = 12), Korean (N = 9), Vietnamese (N = 7), Thai (N = 7), Indonesian (N = 3), Spanish (N = 2), French (N = 1),

achieve this at the English Language Centre by either being enrolled in an ELBP course or attending other different training courses according to their scores and sitting an IELTS test later with a satisfactory result.

²⁹ The actual number of subjects should be 58 because five of them participated in both rounds, hence the number 63.

Italian (N = 1) and Bangla (N = 1). The average length of English study (mainly in their own countries) was 10.5 years and the average length of their stay in Australia was 3.3 months. Their English proficiency levels were indicated by the types of courses they were enrolled in, which were in turn decided by their scores of the placement test or by their IELTS scores, if any. **Appendix II** shows the profile of the recruited subjects in detail.

3.1.3. Difficulties Encountered in the Sampling Procedure

Although the data collection was satisfactory on the whole, there are still a number of issues that need to be addressed in terms of difficulties encountered in the sampling procedure. First, due to the relatively small sampling population and the ethical prerequisite for sampling potential subjects on a volunteer-to-participate basis, random sampling was deemed not practically applicable in the case of the present research. Adopted instead was the convenient stratified sampling in terms of different levels of English proficiency of the students. However, the number of students attending different courses at the centre varied greatly, which posed a problem to the representativeness of the expected different-level sample among the potential population.³⁰ In view of the potential sampling population in the first round of data collection, it was decided that GE 1 and GE 6 were excluded for lack or insufficiency of students; students from GE 2, GE 3B, GE 4B, GE 4C and GE 5 were sampled for the targeted levels.³¹ Contrary to the expected sampling outcome, only two out of 14 students at GE 2, and two out of 15 at GE 3 were recruited as subjects; while 14 out of 30 students at GE 4 and six out of 15 at GE 5 were recruited. In review of the unsatisfactory result of the first-round data collection, it was seen that the unbalanced sample level was mainly due to fewer volunteers among students of lower proficiency levels no matter how hard were the efforts made to recruit them. Those students were more reluctant than those enrolled at higher levels to participate in the project possibly for lack of confidence in their English proficiency and

³⁰ At the time of the first round data collection, there was no class at the beginner level and one preliminary-level class (GE 2; N = 14), three pre-intermediate-level classes (GE 3A-C), four intermediate-level classes (GE 4A-D), one upper-intermediate-level class (GE 5; N = 15), and one advanced-level class (GE 6; N = 11). ETP classes were not sampled in the first round because the focus then was on students enrolled in GE courses.

³¹ The average score for GE 3B is a bit higher than that for GE 3C and a bit lower than that for GE 3A, and is therefore more representative of the level of GE 3. So is the same case for GE 4B and GE 4C, which are more representative of the level of GE 4.

lack of pressure for sitting an IELTS test in the near future, hence less interest. A decision was made in the second round of data collection to shift the focus to those enrolled at higher levels so as to be able to recruit a reasonably well representative sample at the intermediate, upper-intermediate and advanced levels. This proved to be effective: apart from the 4 subjects enrolled at the preliminary and pre-intermediate levels in the first round, the final number of students recruited at higher levels was 18 (intermediate), 21 (upper-intermediate) and 20 (advanced) respectively.³²

Another issue regarding sampling is the recruitment of the first-round participants in the second round of data collection. This turned out to be difficult for the following reasons. A) The planned six-month interval was a bit too long for recruiting some first-round participants; for instance, all the first-round Japanese participants had finished their six-month English training and returned to their own universities in Japan and some had passed the IELTS test score for university entry and left the centre. B) Despite great efforts made to recruit the remaining first-round participants, some were still reluctant to take part perhaps due to lack of interest as well as some other personal reasons. As a result, only 5 out of the 25 first-round participants were recruited in the second round.

Despite all the difficulties encountered in the sampling procedure for the study, every effort was made to resolve them and the sampling outcome was considered to be satisfactory, albeit not ideal. Moreover, it was a valuable first-hand research experience for the researcher.

3.2. The Tasks

This section gives a detailed account of the tasks for data collection including a conversation, elicited repetition, sentence combination, grammaticality judgment (see

³² The variable of the English proficiency levels of the subjects was used in the examination of the results in different tasks of the study. Conveniently, the subjects were classified into three proficiency level groups for analysis: the low level group (including 2 subjects from the preliminary level, 2 from the pre-intermediate level and 18 from the intermediate level), the mid-level group (including 21 subjects from the upper-intermediate level), and the high-level group (including 20 subjects from the advanced level). Note therefore that the proficiency levels thus classified for analysis (low, mid and high) do not match precisely the levels at which the subjects were enrolled, nor do they indicate accurately the level differences that the names of these classified levels might suggest.

Appendix III) and free essay collection. The whole design of the tasks aimed to elicit a variety of styles in interlanguage data in terms of Tarone's continuum of interlanguage styles: Vernacular – Style 2 – Style 3 – Careful Style (Tarone 1983: 152). Along the continuum, the conversation was designed to elicit data in Vernacular Style, the tasks of elicited repetition and sentence combination were for data in Style 2 and the task of grammaticality judgment for data in Style 3. It was hoped that interlanguage data of various styles would provide a multi-dimensional examination of typological universals examined in the study and a more complete picture of interlanguage features of the subjects under investigation. What follows is a description of each of the tasks in the study with respect to its design, purpose, administration, the kind of data sought, and issues regarding reliability and validity.

3.2.1. Conversation – Interview and Role-Play with Cue Card

The conversation comprising an interview and a role-play with a cue card, was designed to elicit spontaneous English interlanguage speech from L2 learners in a simulated natural setting similar to sitting an IELTS speaking test. The design of the conversation resembled the speaking test in the IELTS except that the cue card role-play in the former, unlike that of the IELTS (inserted in the interview), was designed with specific emphasis on elicitation of questions and was conducted after the interview. The interview was aimed to get background information of the subjects such as their country of origin, first language, age, education, years of learning English and length of stay in Australia. The cue card role-play, on the other hand, was specifically targeted for eliciting questions and testing the typological universal regarding question formation (i.e. Universal 11 in Greenberg 1963: 65) in comparison with findings from another study (Eckman et al 1989). In addition, both the interview and the role-play were used to establish a database of naturally-occurring English interlanguage against which some typological universals were tested and syntactic features of L2 learners' interlanguage were investigated.

The interview was informal and less structured as compared with the tasks of elicited repetition, sentence combination, grammaticality judgment, and even the role-play used in data collection. In fact, apart from a few pre-planned questions regarding the subjects'

background information, the interview was quite open in that the subjects were free to express their ideas and to elaborate on what interested them. Moreover, those personal questions were not asked in a fixed order but intermingled with other questions of general topics in a natural way. The interview in this sense was more like a natural conversation which focused on meaning/communication rather than on form, and data thus gathered was similar to the language produced in a natural setting. However, due to the nature of the open interview, it was hard to obtain enough relevant data against which to test typological universals such as the Accessibility Hierarchy though such data were examined in this respect. Consequently, the data from the interview were mainly examined descriptively and analysed qualitatively to explore common interlanguage features such as in tense/aspect of the verb and singularity/plurality of the noun, and some other prominent features. In view of all this, the interview in this study, unlike the above-mentioned tasks, was characteristic of a synthetic approach, heuristic objective, very low degree of control/manipulation over the research context and of explicitness of data collection procedure.

The role-play, though communicatively oriented, also directed the subjects' attention towards form to the extent that in either of the two cue cards, *yes/no* and *wh* questions were explicitly required in performing the task. This might have some impact on the questions the subjects produced, that is, they might pay more attention to correctly forming questions required in this task context than to those asked more naturally (though less likely) in the interview. Furthermore, they had less freedom of elaboration than in the interview due to the essential information-seeking role they had to fulfil as a television buyer or a tourist with reference to the cues provided. Nevertheless, the communicatively-oriented role-play with certain degree of focus on form (questions) would increase the number of instances of questions in the data against which the typological universal of question formation was to be tested. In so doing, issues regarding reliability and validity were carefully addressed. As compared with the interview, this task was typical of an analytic approach, deductive objective, moderate degree in both control/manipulation over the research context and in explicitness of data collection procedure. The research question for this task was: Does the typological universal –

Greenberg's Universal 11 regarding question formation (Greenberg 1963) – make the same predictions for L2 learners' production of English in the task of role-play? The research question was converted into the following specific prediction hypothesis:

Hypothesis 1: In the conversation task, the questions produced by individual L2 learners are predicted to follow the implicational hierarchy: *Yes/no* inversion implies *wh* inversion, which in turn implies *wh* fronting. In terms of quantitative prediction for produced questions in the task, the percentages for instances of *yes/no* inversion should not be higher than those for instances of *wh* inversion, which should in turn not be higher than those for instances of *wh* fronting, regardless of the learners' English proficiency levels and first language backgrounds.

In this hypothesis, the predictor – the implicational hierarchy (or the two implicational universals) concerning question formation, is the independent variable, while the subjects' performances on the task – actual questions produced by them become the (predicted) dependent variable. The independent variable was assumed to predict the dependent variable via a number of operational steps proposed by Eckman et al (1989). A similar procedure to that of Eckman et al (1989) was followed to make this study comparable to theirs. Based on operational definitions of some grammatical constructions such as *wh* questions, *yes/no* questions and inversion, relevant utterances were classified and tied to those grammatical constructions.³³ The grammatical constructions were then tied to one of the three general features of word order – Features A, B and C.³⁴ Finally, the two implicational universals can be justified or falsified in the interlanguage data of this study by testing whether Feature C implies Feature B, which in turn implies Feature

³³ Grammatical constructions which are identified in this study are identical with those defined in Eckman et al (1989: 201-4) with some modifications where necessary. For example, "question pronouns" (Eckman et al 1989: 175) and "*Wh* pronoun" (180) are avoided, and "question words or phrases" (Greenberg, 1963: 83) is consistently used instead. The latter seems more accurate than the former because *Wh* questions include questions formed not only with question pronouns such as *who*, *whom*, *what*, but with question adverbs and question pronoun/adverb phrases as well such as *where*, *when*, *why*, *how*, *how much* and *how long*. "Why did Sue stop drinking?" (Eckman et al 1989: 182) is an example in point. Semantic criterion is also used in discriminating those applicable "*What is reading Mary now?*-type sentences" (Eckman et al 1989:202) from those inapplicable questions in which question words function as subject such as "What/Who is bothering Mary now?".

³⁴ Features A, B and C are the equivalents of Patterns A, B and C in Eckman et al (1989).

A. In other words, the implicational universals will be justified if they are testified in the data in quantitative terms and they will be falsified if proved otherwise.

As compared with Eckman et al (1989), this study examined more subjects (50 versus 14) of more diverse language-speaking backgrounds (10 versus 3). It also involved further finer examination with controlled variables such as proficiency level and first language. Eckman et al (1989) claimed that their task of eliciting questions around story squares reflects a content-conscious, communicative language variety that is underlain by the same type of unconscious knowledge in which primary languages are rooted, and that thus elicited data stand the best chance to reveal *acquired* rather than *learned* knowledge (cf. Krashen 1977, 1980) of interlanguage learners. In comparison, the present study uses a different task – a natural conversation comprising an interview and a role-play, which should elicit the same type of *acquired* knowledge of interlanguage learners because the task reflects a less content-conscious and more communicative interlanguage variety. Eckman et al's more content-centred elicitation of questions, however, elicits more questions from the 14 subjects than does the task of the present study from 52 subjects. Despite these differences, this study, in essence, replicates Eckman et al (1989) in that it aims to examine the validity of the same implicational universals regarding constituent order of questions in spontaneous speech of L2 learners and adopts the same conceptual framework and methodological procedure used in their study.

Apart from independent and dependent variables that were the focus of all the tasks in this study, there were two sets of variables – “subject and extraneous variables” (Seliger and Shohamy 1989: 92) that needed to be controlled and manipulated in order to enhance the reliability and validity of the tasks. The former related to such factors as age, sex, native language, years of learning English, length of stay in Australia, and English proficiency, and the latter involved variables such as the native speaker or the researcher, the administrator of the task, and issues concerning the design of the task.

In terms of subject variables, those of age, years of learning English, and length of stay in Australia were not greatly differentiated, and were thus treated roughly as homogeneous

groups in their respective terms. For example, the average age for all the subjects was 23.8 years with only four over the age of 30, and the average length of their stay in Australia was 3.3 months with only two having 8 and 12 months respectively. The average length of their learning English was 10.5 years with four having over 15 years and another four (all are Vietnamese-speaking subjects) less than 5 years, and they all shared similar English-learning experience – formal classroom instruction in their home countries. The variable ‘English proficiency’ was controlled in recruiting potential subjects, and was used, together with the other subject variable (first language), as a moderator variable in this study. That is, the two subject variables were examined throughout the tasks in this study to see whether they would ‘moderate’ or affect overall results obtained from examining the relationship between the independent variable (typological universals) and the dependent variable (the subjects’ responses in different tasks).

The extraneous variables for this task such as the native speaker, the administrator of the task, and issues concerning the design and administration of the task were taken into consideration and controlled properly in the data collection procedure. Measures were taken to ensure the effectiveness of the interview and the role-play and to enhance their reliability and validity. First, the design of the conversation was made comparable to that of the speaking band of an IELTS test but was less formal than the latter because the conversation was less structured and more personalised and conducted in an informal situation, allowing the subject to elaborate with free responses. The subjects were familiar with the task situation and would regard their participation as a practice opportunity for a real test; as a result, undesirable effects such as the Hawthorn effect³⁵ could be reduced to minimum. Moreover, the topics covered in the conversation, in addition to their similarities to the IELTS test, were also made as real and relevant to their life and study as possible, hence more communicatively oriented. For example, topics in the interview covered not only some background information about the

³⁵ Richards et al (1992: 165-6) defines *Hawthorn effect* as “(in research) the effect produced by the introduction of a new element into a learning situation. For example, if a new teaching method is used, there may be an improvement in learning which is due not to the method, but to the fact that it is new. Later on, the improvement may disappear.”

participants, but their study, life, hobbies, sports, and interesting experiences as an international student in Australia, and those of the role-play also focused on their life-related matter (buying a secondhand television and visiting Melbourne). In this way, the subjects' naturalistic interlanguage rather than their metalinguistic linguistic knowledge of English was expected to occur.

Next, a pilot study was conducted before the real data collection in order to try out the planned data collection procedure including the time allotted to each task, the test items and using test instruments. In the pilot test some problems arose concerning wording and arrangement of test items and ways of administering each task, which were resolved accordingly before the commencement of data collection. Consequently, the revised tasks and procedure were implemented in the study.

The issue concerning recruitment and training of the native speaker as the administrator of the conversation was also given considerable attention. A suitable native English speaker was recruited, among many candidates, on the basis of her life and work experiences and personality. In order to make her task consistent and accurate across all participants, she was trained on several occasions including the pilot study. She was well acquainted with the format of the interview including the gambits for beginning, shifting and ending conversations, topics to cover, techniques for making subjects feel at ease and relaxed to talk, tempo of the speech, control of theme and time, and even monitoring of the tape-recorder. In effect, the native English speaker proved to be successful in administering the conversation, particularly in motivating the subjects in a meaningful, vivid and life-like communication with her conversational skills.

Finally, the internal and external validity for the heuristically-examined conversation was taken into account in both data collection procedure and interpretation of the data. As has been illustrated above, the data collected in the conversation were representative of naturalistic data. The data were audio-recorded and kept for possible re-examination in the future. The transcripts of the data were double-checked and the interpretation of the

data was re-inspected in line with relevant current theories in an effort to minimise the effect of subjectivity and personal bias on the part of the researcher.

3.2.2. Elicited Repetition

The task of elicited repetition was designed to test the Accessibility Hierarchy in comprehension data of L2 learners' interlanguage. More specifically, the research question for this task was: Does the typological universal – the Accessibility Hierarchy – make the same predictions for L2 learners' comprehension of English in the task of elicited repetition? This research question was then converted into the following prediction hypothesis:

Hypothesis II: In the task of elicited repetition of English restrictive relative clauses, L2 learners' correct instances of repetition on SU position will be greater than or as many as those on DO position, which will, in turn, be greater than or as many as those on IO position, and so on for the other lower positions on the Accessibility Hierarchy, regardless of the learners' English proficiency levels and first language backgrounds.

In this hypothesis, the independent variable is the Accessibility Hierarchy, and the dependent variable is the subjects' performances on the task – actual instances of repetition produced by them. That is, the independent variable that the degree of accessibility varies in accordance with different relativised positions on the hierarchy among the world's languages, is assumed to predict the same pattern in the dependent variable – interlanguage performances in terms of instances of correct repetition on each relativised position. Unlike the role-play, this task has a high degree in both control/manipulation over the research context and in explicitness of data collection procedure with relative focus on form in "Style 2" rather than on meaning in the "Vernacular Style" (Tarone 1983). Similar to the role-play, the task is characteristic of an analytic approach and deductive objective, and has the same focus on relation between independent and dependent variables with the same set of controlled subject variables.

Other variables that could affect Hypothesis II were controlled in the data collection procedure of the task. First, the items of the test were designed to the effect that they tapped and measured the subjects' comprehension of relativised grammatical positions in relative clauses rather than other factors such as their literacy level or their memory capacity. So factors such as semantic content of words, number of syllables in words and length of sentences were taken into account in designing the items of the task. To lessen the impact the subjects' varying English literacy capacity might have on their performances, comparatively easy and common English words were used in test items. To minimise the impact the subjects' varying memory capacity might have on their performances, relatively short sentences were chosen as test items with each one containing a relative clause and having 9 or 10 words ranging between ten to thirteen syllables in all. To increase reliability of the task, each of the six relativised positions on the Accessibility Hierarchy was represented by two items totaling twelve test items in the task. All the twelve relative clauses had the same grammatical head noun, the subject in the matrix clause, and the six types of relative clauses were mixed up in the presentation sequence.

To increase the validity of the task, the quality of the recorded sentences was also attended to. A native English speaker with expertise in English phonetics and phonology was invited to do the recording of the sentences. In the sound lab of the Department of Linguistics at Monash University, after several rehearsals the expert read out the sentences one by one at a normal pace with a 15-second pause between sentences. These sentences were recorded onto a disc, which was then technically cleared of any unwanted noise by a technician. The voice of the speaker was very clear and the quality of the recording was very good.

Finally, the effect of the researcher on data collection of the repetition task was addressed seriously as well. After finishing the conversation with the native English speaker, the subject was asked to proceed to the researcher in the opposite corner of the classroom to complete the repetition task. To administer the task consistently, the researcher followed the same procedure: a brief and clear instruction was given to each and every subject

before the task to ensure that they understood the requirement for the task. Then they were encouraged to try to repeat after the recorded sentences as much as possible and their repetition was taped-recorded.

3.2.3. Sentence Combination

The task of sentence combination was also designed to test the Accessibility Hierarchy in L2 learners' interlanguage production. The research question for this task was: Does the typological universal – the Accessibility Hierarchy – make the same predictions for L2 learners' production of English in the task of sentence combination? This research question was converted into the following prediction hypothesis:

Hypothesis III: In the task of sentence combination involving English restrictive relative clauses, L2 learners' correct instances of combining sentences on SU position will be greater than or as many as those on DO position, which will, in turn, be greater than or as many as those on IO position, and so on for the other lower positions on the Accessibility Hierarchy, regardless of the learners' English proficiency levels and first language backgrounds.

In this hypothesis, the independent variable, again, is the Accessibility Hierarchy, and the dependent variable is the subjects' performances on the task, actual instances of combined sentences. The independent variable, again, is assumed to predict its corresponding manifestation in the dependent variable: the degree of accessibility of grammatical positions on the hierarchy should correlate quantitatively with instances of predicted sentence combination for each relativised position in the task. Compatible with the task of elicited repetition, this task is also characteristic of an analytic approach and deductive objective, and has a high degree in both control/manipulation over the research context and in explicitness of data collection procedure. Unlike the task of elicited repetition, however, data gathered from this task should be ranked more towards the "Careful" end of the style continuum (Tarone 1983) as "Style 3" instead of "Style 2" because this task involved metalinguistic knowledge as well as relative focus on form.

Apart from the controlled subject variables of sex, first language and English proficiency, extraneous variables concerning this task were controlled as well in designing and administering the task. There were 14 pairs of sentences in the task and the task required the subjects to combine each pair into a sentence with a relative clause. Similar to the repetition task, each of the six relativised positions on the Accessibility Hierarchy was, predictably, realised in two instances except for the position of GEN, which was divided into GEN-SU and GEN-DO with each having two instances of predicted combination. Different from the design of elicited repetition, however, is that for the same relativised position, the two instances had different grammatical head nouns in the matrix clause, one is the subject and the other the direct object. This was designed to test more finely whether the position of the head noun in the matrix clause might have any effect on the same relativised position in the relative clause. The fourteen pairs of sentences were arranged randomly.

To ensure that this task measured the subjects' ability to combine sentences in relation to different relativised positions rather than something else, unfamiliar words that might hinder the subjects' understanding of sentence meaning were avoided in working out task items. Care was taken in administering the task to the subjects. The task was administered by the researcher to the subjects immediately after the task of elicited repetition and was expected to be fulfilled within 10 minutes. The subjects were given the task paper, briefly instructed on how to combine a pair of sentences into a single one, and then reminded of the illustrative example, that is, the second sentence in a pair became a relative clause. Whatever they wrote in combining sentences constituted raw data for this task.

Properly designed and administered, sentence combination should be an ideal test tool in studies of relative clause formation in SLA because of its high level of reliability and validity. In a recent study which focused on cross-sectional restrictive relative clause acquisition and relative clause test types among Japanese students, Ito (2000) examined the reliability and validity of different types of tests and the difference in results caused by application of different test types. Among the four types of relative clause tests used in the study – translation, cloze procedure, grammaticality judgment, and sentence

combination – Ito finds that sentence combination shows high reliability in internal consistency and the highest validity.

3.2.4. Grammaticality Judgment

The task of grammaticality judgment was designed to test comprehension of relative clauses as well as a number of typological universals particularly those motivated by Hawkins (1999) in comprehension data of L2 learners' interlanguage. In the task were 36 stimulus sentences (statements or questions), each of which was followed by three judgment alternatives – *acceptable English*, *unacceptable English*, and *not sure*. Related to relative clauses were S1-S8, S11 and S19 in which five were formed correctly and the other five contained five types of non-standard RC forms including resumptive pronoun, incorrect RC marker, omission of obligatory relative marker, non-adjacency and prenominal relative clause. Related to the position of interrogative pronouns were three stimulus sentences in which interrogative pronouns were fronted (S9), placed in situ (S12), and inserted in the middle (S15). The rest of the 36 stimulus sentences were respectively targeted for examining implicational hierarchies/hypotheses such as the Clause Embedding Hierarchy, the Subordinate Gap/No Gap Hierarchy, the Hierarchy for Complementisers, the Bridging Verb Hierarchy, the Head Noun Phrase Hierarchy, and the Valency Completeness Hypothesis.³⁶ The overall research question for the task regarding these implicational hierarchies/hypotheses was: Do these implicational hierarchies/hypotheses make the same predictions for L2 learners' comprehension of English in the task of grammaticality judgment? This research question is converted into a number of null hypotheses, which are expressed below respectively.³⁷

Hypothesis IV: There is no relation between the filler-gap domains in different types of clauses as noted in the Clause Embedding Hierarchy (Hawkins 1999), and the ESL

³⁶ Refer to **Appendix 1** for an overview of the typological universals tested in the interlanguage data in this study. Refer to Chapter 2 (2.1.3) for a detailed explanation of these hierarchies/hypotheses. Refer to chapter 4 (4.3) for the test results, detailed analysis and discussion.

³⁷ The statistical procedure of Chi-square (see Hatch and Lazaraton 1991: 393-424) was used to test *Hypotheses IV-IX* to see whether they could be rejected. The significance level was set at .05 for all the six null hypotheses. If the null hypothesis was rejected (i.e. $p < .05$), the related universal would be confirmed in the second language data (specifically, in the grammaticality judgment task); if otherwise, the related universal would be disconfirmed.

learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing a filler-gap domain with the gap in a infinitival phrase will not be greater than those of judging as acceptable the sentence containing a filler-gap domain with the gap in a finite subordinate clause, which, in turn, will not be greater than those of judging as acceptable the sentence containing a filler-gap domain with the gap in a complex NP.

Hypothesis V: There is no relation between the filler-gap domains containing an extra gap and no extra gap in a subordinate clause as noted in the Subordinate Gap/No Gap Hierarchy (Hawkins 1999), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing a filler-gap domain in which a matrix filler can be matched with a gap in a subordinate clause of complexity *n* containing another gap, will be greater than those of judging as acceptable the sentence containing a filler-gap domain in which the matrix filler can be matched with a gap in subordinate clause of complexity *n* containing no other gap.

Hypothesis VI: There is no relation between the filler-gap domains containing different types of complementisers as noted in the Hierarchy for Complementisers (Kluender 1992), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing *that* as acceptable will not be greater than those of judging as acceptable the sentence containing *if*, which, in turn, will not be greater than those of judging as acceptable the sentence containing *whether*.

Hypothesis VII: There is no relation between the filler-gap domains containing bridging verbs of varying semantic specificity as noted in the Bridging Verb Hierarchy (Hawkins 1999), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing a bridging verb (complex) *V* of semantic specificity within a filler-gap domain, will be not be greater than those of judging as acceptable the

sentence containing a bridging verb (verb complex) V' with more semantic specificity than V.

Hypothesis VIII: There is no relation between the filler-gap domains containing NPs with head nouns of varying semantic specificity as noted in the Head Noun Phrase Hierarchy (Hawkins 1999), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing an NP with a head noun (phrase) N of semantic specificity within a filler-gap domain, will not be greater than those of judging as acceptable the sentence containing an NP with a head noun (phrase) N' of more semantic specificity than N.

Hypothesis IX: There is no relation between filler-gap domains that are valency complete or valency incomplete as noted in the Valency Completeness Hypothesis (Hawkins 1999), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing a filler-gap domain that includes the subcategorisers for all phrases within the domain that contains the gap, will not be greater than those of judging as acceptable the sentence containing a filler-gap domain that does not include all the subcategorisers for all phrases within the domain that contains the gap.

In each of the above six hypotheses, the independent variable is an implicational hierarchy or hypothesis which comprises two or more levels, and the dependent variable is the subjects' actual judgment choices measured in terms of frequency tally. Again, the independent variable is assumed to predict the pattern for the dependent variable and the same set of subject variables were controlled and examined to see whether they might contribute to any variation in the relation between independent and dependent variables. Compared with all the other tasks, this task is the most typical of an analytic approach and deductive objective, and has a very high degree both in control/manipulation over the research context and in explicitness of data collection procedure. In view of Tarone's

style continuum, the responses elicited from this task should definitely be ranked as "Style 3" with a strong focus on both form of language and metalinguistic ability of the subjects.

Most of the task items concerning testing the typological universals and hierarchies were quoted from Hawkins (1999) in an attempt to directly test the validity of those claims associated with his motivated universals in the L2 learners' comprehension data of the present study. Each of the typological universals was tested by a group of sentences. However, the sentences of the same group were scrambled with those of other groups in sequencing test items so as to avoid an ad hoc judgment resulting from grouping similar sentences together. For instance, the Head Noun Phrase Hierarchy was tested in three stimulus sentences numbered S14, S17, and S23 respectively rather than grouped together. It was hoped that this arrangement would better tap the subjects' intuitive on-line judgment without being affected by immediate reference inherent in the grouping of similar sentences together.

In the task the subjects were told to make a judgment of each stimulus sentence by circling or ticking only one of the three choices – *acceptable*, *unacceptable* and *not sure*.³⁸ The subjects were told to complete the task within ten minutes and not to spend too much time on some difficult ones. If they exceeded the time limit, the researcher would kindly remind them of the time.

The task of grammaticality judgment has been widely used in SLA to test L2 learners' metalinguistic ability which is believed to indicate interlanguage competence. For example, in a study which examined the reliability of L2 grammaticality judgments among Chinese-, Korean- and Japanese-speaking English learners, Gass (1994) found that reliability issues are related to a L2 learner's indeterminacy and incomplete

³⁸ Following the common practice in a grammaticality judgment task, I used *acceptable* and *unacceptable* rather than *grammatical* and *ungrammatical* as possible choice alternatives in this task in order to avoid the subjects' total reliance on their knowledge of prescriptive grammar of English (see Munnich et al 1994: 229). Furthermore, I also included another possible choice alternative *not sure* indicating the status of indeterminacy in the grammatical knowledge of the subjects, so as to tap their complete intuitive

grammatical knowledge, and that grammaticality judgments can indeed reflect the patterns of L2 learners' grammatical knowledge of the target language.

3.2.5. Free Essay Writing

Free essay writing was not a task performed by the subjects in the test venue, but was requested by the researcher of the subjects. In order to gain a comprehensive interlanguage database of the subjects, apart from the data gained from the tasks mentioned above, the researcher asked each subject to provide a copy of one of the essays written in class as a supplement to the data collected. Although they had agreed in their informed consent forms for the researcher to have access to their writing and they were asked earnestly again at the test venue and were provided with a stamped return envelope and 20 cents for copying, many of them still failed to do so. In the end, only 45 essays were collected. Because less control could be exerted over the collection of the subjects' writing, the essays collected vary with respect to topic, genre, length and depth.

These writings generally represented a "Careful Style" (Tarone 1983) in which the subjects produced language forms far more accurately and correctly than in their "Vernacular Style". They were used to test against typological universals such as the Accessibility Hierarchy, and were studied regarding interlanguage features in comparison with the subjects' oral English interlanguage.

3.3. Data Processing

As mentioned before, the whole procedure for data collection took about seven months. Data processing which involved the process of transcribing, coding, tabulating and sorting out data lasted several months. The researcher personally did all the transcribing work and double-checked it all to ensure that all the details transcribed were true to the original recorded material. The transcription for the interview and the task of cued elicitation of questions formed a database of over ten thousand words and that for the task of elicited repetition was about seven thousand words. The background information about

knowledge of the English grammar instead of the traditional way of giving the subject dichotomous preference alternatives, of one or the other of which he/she has to make a choice.

the subjects was obtained mainly through the transcribed database and some missing information was retrieved indirectly through personal communication with some subjects later on. This information was then coded, tabulated and summarised in **Appendix II**.

Raw data in different tasks were classified accordingly, calculated when necessary, and tabulated into various tables ready for analysis. For example, in the task of role-play, all the questions in the conversation for each individual were first tabulated and classified into the categories of *yes/no* questions and *wh*-questions. The *yes/no* questions were then classified into those with inversion and those without, and all the instances of *wh*-questions fell into the feature of *Wh Fronting* with no exceptions. In the case of *wh*-questions, a further classification between those with inversion and those without was made. Then all the instances of these features were calculated in terms of number and percentage, and were tabulated in line with Eckman et al (1989) so as to make the results of the study comparable to their findings.

In the repetition task, the transcribed data were first classified as belonging to different features such as correct repetition, near correct repetition, partial or non-repetition, omission of obligatory RC marker, incorrect RC marker, different meaning, different relativisation and topic-comment structure. Thus classified data were calculated in terms of the number of actual instances for each feature and the percentage each feature had over all instances, and then were expressed in tabulated forms. The same method was applied to the sentence combination task and the naturalistic data as well.

With respect to data analysis, descriptive statistics such as percentages were constantly employed to describe relative proportions and tendencies, and at times referential statistics such as the Chi-square test was used to examine some of the nominal data to see whether the relation between variables would be statistically significant. Apart from quantitative analysis of the data, qualitative analysis was also applied to the data in relation to some specific methods as well as the findings in the relevant literature. The detailed analysis of the results and related discussion is, however, the focus of chapter four, to which we now turn.

Chapter Four Results and Analysis of Findings

This chapter consists of four sections dealing with the results and analysis of findings concerning both analytically and deductively investigated data (sections 4.1-4.3) and holistically and heuristically examined data (section 4.4) in the study. Section one (4.1) gives an analysis and discussion of results regarding implicational universals of *yes/no* and *wh* questions (Greenberg 1963). Section two (4.2) provides an analysis and discussion of results regarding the Accessibility Hierarchy (Keenan and Comrie 1977). Section three (4.3) includes an analysis and discussion of results regarding implicational hierarchies/hypotheses such as Hierarchy for Complementisers (Kluender 1992), Clause Embedding Hierarchy, Subordinate Gap/No Gap Hierarchy, Bridging Verb Hierarchy, Head Noun Phrase Hierarchy, and Valency Completeness Hypothesis (Hawkins 1999). Section four (4.4) is a further examination of interlanguage features in the naturalistic data of the study with reference to the results of the other three sections.

4.1. Results regarding Implicational Universals of *Yes/No* and *Wh* Questions

In this section, the results of the subjects' performances in the conversation data relating to the implicational universals of *yes/no* and *wh* questions (Greenberg 1963) are presented in 4.1.1, then interpreted and discussed in 4.1.2, and finally followed with summarising remarks in 4.1.3.

4.1.1. Results of Data

Recall (in 3.2.1) that Greenberg's (1963) implicational universals regarding the word order of questions were converted into a quantitative prediction hypothesis, which is restated in (1).

- (1) **Hypothesis I:** In the conversation task, the questions produced by individual L2 learners are predicted to follow the implicational hierarchy: *Yes/no* inversion implies *wh* inversion, which in turn implies *wh* fronting. In terms of quantitative prediction

for produced questions in the task, the percentages for instances of *yes/no* inversion should not be higher than those for instances of *wh* inversion, which should in turn not be higher than those for instances of *wh* fronting, regardless of the learners' English proficiency levels and first language backgrounds.

Accordingly, all the questions in the conversation were sorted out from raw data, then they were classified into two groups of utterances: those that exemplified one or the other of the three basic order features that the universal make reference to and those did not exemplify them. Specifically, the total number of *wh* questions is 166 of which 153 instances involve inversion. The total number of *yes/no* questions is 349 of which 313 are applicable.³⁹ Out of the 314 *yes/no* questions, 241 instances involve inversion. The performances of all the subjects in terms of instances manifested in the three order features are presented in **Appendix IV**. The following tables (4.1.1-4.1.3) provide the resulting counts as relevant for the implicational universals for different proficiency groups.⁴⁰

³⁹ Excluded from the study were eight subjects who produced 36 *yes/no* questions in all but did not produce any *wh* questions at all in the task of conversation, hence inapplicability for testing the universal at issue.

⁴⁰ Table 4.1.1 provides utterance counts for 16 subjects in the low-level group including a second-round re-participant, S1, Table 4.1.2 for 18 subjects in the mid-level group, and Table 4.1.3 for another 18 subjects in the high-level group including a second-round re-participant (M7), who, at the time of second-round data collection, had just finished an English training course at the English language centre, passed the IELTS test, received a full offer for a master course at Monash University. Note that the level differences might not be as great as the names of these level groups suggest (see 3.1.3 for details). The tables employ the same set of abbreviations from Eckman et al (1989: 176) for presentation:

" VS = verb-before-subject order (Inversion);
#*wh* = initial positioning of *wh* words (*Wh* Fronting);
YNQ = *yes/no* questions; and
WHQ = *wh* questions."

Table 4.1.1 Token Counts and Percentages for
Yes/no Inversion, Wh Inversion and Wh Fronting (low level group)

1	2	3	4	5	6	7	8	9
	Yes/no Questions			Wh Questions			Wh Questions	
Subject	Total YNQ	VS in YNQ No.	%	Total WHQ	VS in WHQ No.	%	#wh in WHQ No.	%
S1a	8	4	50	3	2	67	3	100
K2	2	1	50	2	2	100	2	100
J10	6	3	50	1	1	100	1	100
M6	6	4	67	1	1	100	1	100
J11	3	2	67	1	1	100	1	100
J4	7	5	71	4	4	100	4	100
M4	5	4	80	3	3	100	3	100
M1	6	5	83	3	3	100	3	100
J2	6	5	83	2	2	100	2	100
S1b	6	5	83	4	4	100	4	100
K3	7	6	86	3	3	100	3	100
K6	7	6	86	3	3	100	3	100
V1a	4	4	100	4	4	100	4	100
J3	6	6	100	5	4	80	5	100
J7	2	2	100	2	2	100	2	100
M8	2	2	100	3	3	100	3	100
TOTAL	83	64	77	44	42	95	44	100

Table 4.1.2 Token Counts and Percentages for
Yes/no Inversion, Wh Inversion and Wh Fronting (mid level group)

1	2	3	4	5	6	7	8	9
	Yes/no Questions			Wh Questions			Wh Questions	
Subject	Total YNQ	VS in YNQ No.	%	Total WHQ	VS in WHQ No.	%	#wh in WHQ No.	%
M15	7	3	43	3	3	100	3	100
M2a	12	6	50	4	2	50	4	100
S2	2	1	50	3	2	67	3	100
T7	10	5	50	2	2	100	2	100
K1	9	5	56	9	8	89	9	100
T2	9	5	56	3	3	100	3	100
In1	7	4	57	2	1	50	2	100
M3	6	4	67	2	2	100	2	100
M10	7	5	71	3	3	100	3	100
In2	8	6	75	4	4	100	4	100
K8	9	7	78	2	2	100	2	100
K5	5	4	80	2	2	100	2	100
T6	5	4	80	4	4	100	4	100
J8	14	13	93	4	4	100	4	100
K4	4	4	100	4	4	100	4	100
M12	3	3	100	2	2	100	2	100
M13	17	17	100	2	2	100	2	100
M14	3	3	100	4	4	100	4	100
TOTAL	137	99	72	59	54	92	59	100

Table 4.1.3 Token Counts and Percentages for
Yes/no Inversion, *Wh* Inversion and *Wh* Fronting (high level group)

1	2	3	4	5	6	7	8	9
Subject	<u>Yes/no Questions</u>			<u>Wh Questions</u>			<u>Wh Questions</u>	
	Total	VS in YNQ		Total	VS in WHQ		#wh in WHQ	
	YNQ	No.	%	WHQ	No.	%	No.	%
M9	4	1	25	5	5	100	5	100
T5	7	4	57	3	2	67	3	100
In3	6	4	67	6	6	100	6	100
T4	3	2	67	2	2	100	2	100
T1	8	6	75	3	3	100	3	100
M7a	6	5	83	3	1	33	3	100
V6	6	5	83	1	1	100	1	100
K7	6	5	83	3	3	100	3	100
V7	6	5	83	2	2	100	2	100
M7b	9	8	89	1	1	100	1	100
V4	1	1	100	6	5	83	6	100
V5	8	8	100	3	3	100	3	100
B1	1	1	100	9	8	89	9	100
It1	1	1	100	4	4	100	4	100
M11	7	7	100	1	1	100	1	100
F1	7	7	100	4	4	100	4	100
T3	4	4	100	4	3	75	4	100
M5b	4	4	100	3	3	100	3	100
TOTAL	94	78	83	63	57	90	63	100

Column 1 in all the three tables lists subjects by code names with a capital letter indicating a first language and a digital figure indicating a sequential number assigned to a subject within that particular first language group (see **Appendix II** for details). The order of the subjects follows the order of increasing percentages for the first implicans (i.e. percentages for Feature C – inversion in *yes/no* questions). The last line in each table provides aggregate figures.

Column 2 shows the total number of *yes/no* questions obtained for each subject. Columns 3 and 4 bear on the first implicans of the implicational hierarchy. The former shows the number of those *yes/no* questions in which *Yes/no* Inversion was used and the latter gives the percentages of instances of *Yes/no* Inversion over the total number of *yes/no* questions obtained.

Similarly, Columns 5 provides the total number of *wh* questions obtained for each subject. Columns 6 and 7 bear on Feature B – inversion in *wh* questions, which acts

simultaneously as both the first implicatum and the second implicans of the implicational hierarchy. Column 6 provides the number of those *wh* questions in which *Wh* Inversion was manifest with Column 7 showing the percentages of instances of *Wh* Inversion over the total number of *wh* questions.

Finally, Columns 8 and 9 bear on Feature A – Sentence-initial question words/phrases (*Wh Fronting*), which is the second implicatum of the implicational hierarchy. Again, Column 8 gives the number of those *wh* questions in which *Wh Fronting* was attested with Column 9 providing the percentages of instances of *Wh Fronting* over the total number of *wh* questions.

4.1.2. Interpretation of Data and Discussion

Eckman et al (1989) interpret their data in terms of what they call the Absolute Existence Interpretation. By this approach, they set up a high percentage of 90 to determine the existence of a pattern in interlanguage. If a pattern reaches this percentage or higher, it is marked as plus (+); if the pattern is below this percentage, it is marked as minus (-). For confirmation of an implicational universal such as Pattern C implies Pattern B, the percentages for the two patterns must both reach the 90% threshold with a configuration of “+ +”. A violation of the universal would be a configuration of “+ -”, while the other two possible configurations (“- -” and “- +”) would not be taken to confirm, but would nonetheless be consistent with, the universal. However, they find that an alternative approach called the Relative Existence of Interpretation in terms of quantitative predictions about implicational universals has two advantages over the Absolute Existence Interpretation. The former does not have to decide arbitrarily as to the percentage thresholds, however high they may be, for determining the presence or absence of a grammatical pattern. More significantly, the former makes stronger claims about implicational universals than the latter because it entails the former, but not vice versa.

For interpretation of the results in this study, the Relative Existence Interpretation (Eckman et al 1989) is adopted. That is, to determine whether an implicational universal

holds in the data, it is the comparison of relative percentages rather than the determination of absolute threshold percentages of implicans and implicatum that matters. In other words, where the percentages for the implicatum are higher than or at least as high as those for the implicans, the implicational universal is taken to hold; otherwise it is not.

As can be seen from the tables 4.1.1-4.1.3, Column 9 shows the ceiling effect of results for Feature A (*Wh* Fronting), that is, all subjects fronted the *wh* question word/phrase 100% of the time. This demonstrates that there are no violations for the implicational universal "*Wh* Inversion implies *Wh* Fronting" in terms of comparison of two percentages for Features B (*Wh* Inversion) and A (*Wh* Fronting) because percentages for the former are never greater than those for the latter. Thus in consistence with the same results for this implicational universal in Eckman et al (1989), the data from this study do show support for the implicational universal "*Wh* Inversion implies *Wh* Fronting".

However, when comparing the percentages in Column 4 (for Feature C) and Column 7 (for Feature B), we can find that although most percentages in the latter are either higher or equal to those in the former, there are still some instances in the reverse order (6 out of 52). This shows that, generally, the data lend support to the implicational universal "*Yes/no* Inversion implies *Wh* Inversion", there are nevertheless six counter-examples to the universal. This result also agrees with what Eckman et al (1989) find in their study. Via comparison of relative percentages of instances for the two patterns (i.e. *Yes/no* Inversion and *Wh* Inversion) in each individual subject's data, they find three (out of the 14 subjects) violations, among which only one instance, after further examination, is taken to be a significant violation.

An examination of the six counter-examples was conducted following Eckman et al (1989). These counter-examples are shown in Table 4.1.4.

Table 4.1.4 Counter-examples to "Yes/no Inversion implies *Wh* Inversion"

1 Subject	2 3 4 Yes/no Questions			5 6 7 Wh Questions		
	Total YNQ	VS in YNQ		Total WHQ	VS in WHQ	
		No.	%		No.	%
J3	6	6	100	5	4	80
In1	7	4	57	2	1	50
M7	6	5	83	3	1	33
V4	1	1	100	6	5	83
B1	1	1	100	9	8	89
T3	4	4	100	4	3	75

As mentioned before, because of the nature of the data collection (less controlled and more communicatively-oriented), there are far fewer questions elicited from the subjects in this study than those from those in Eckman et al's study. In view of interlanguage variability (Tarone 1979), one occurrence of a feature can not be taken to show systematicity or violation of a rule and thus all those subjects who produced only one instance in either of the two features are eliminated from analysis. Thus viewed, nine out of the 52 subjects were excluded from analysis including two counter-examples (V4 and B1 in the above table), and among the 43 analysable subjects, four subjects (J3, In1, M7 and T3) are regarded as exceptions to the universal.

However, a further inspection shows that the lower percentages of the implicatum for three subjects (J3, In1 and T3) are caused by just one token, hence not regarded as significant disconfirming instances. Only M7 constitutes a significant violation of the universal that needs to be accounted for. Based on the analysis so far provided, it can be concluded that both universals in the implicational hierarchy are upheld by the data. While the "Yes/no Inversion implies *Wh* Inversion" universal has only one exception (i.e. M7), the "*Wh* Inversion implies *Wh* Fronting" universal holds without exceptions.

The issue about the validity of test interpretation that concerned Eckman et al (1989) is also addressed in this study. Eckman et al (1989) claimed that their task of eliciting questions around story squares reflects a content-conscious, communicative language variety that is underlain by the same type of unconscious knowledge in which primary languages are rooted, and that thus elicited data stand the best chance to reveal *acquired* rather than *learned* knowledge (cf. Krashen 1977, 1980) of interlanguage learners. In this

study, the data collected should reflect a less content-conscious but more communicative language variety, and thus should also stand the best chance to reveal acquired knowledge of interlanguage learners.

In order to determine whether the experimental conditions under which the elicited data were biased in favour of the universal, or specifically, whether the frequency of inversion in *yes/no* questions was depressed in favour of confirmation-seeking type of *yes/no* questions in the task of working out a sequence of events based on pictures,⁴¹ Eckman et al (1989) used the same elicitation procedure to test four native speakers of English and compared their results with those of their interlanguage speakers. They found that the native speakers did invert *yes/no* questions significantly more frequently than the interlanguage speakers, based on which they concluded that the experimental conditions did not bias the results in favour of the universal. In view of the experimental conditions of this study, the task of role-play, in which subjects seek information by initiating questions rather than confirming information through guessing, is less susceptible to bias in favour of uninverted *yes/no* questions. Nevertheless, the effect of the native speaker interviewer in the task of conversation needs examination. Because in the task of conversation the role-play follows the interview in which the interviewer seeks information by asking questions, it is decided to examine whether the way the interviewer asks questions might bear on the subjects' performances on question formation.

Throughout the conversations with the subjects, the native speaker produced 1,724 questions, among which 1,044 are *yes/no* questions. Among these *yes/no* questions, 974 (93.3%) are inverted and 70 (6.7%) are uninverted. On average, the native speaker used 16.2 inverted *yes/no* questions and 1.2 uninverted *yes/no* questions to converse with each of the 60 subjects. These data show that the native speaker did invert far more frequently (93.3%) in *yes/no* questions than did the interlanguage speakers (76.7%). Furthermore, the differences in ranges of *yes/no* inversion are also clear: the range of inversion for the subjects is 25-100%, while the range for the native speaker is 71-100%. In fact, apart

⁴¹ In English, it is quite communicatively appropriate to ask a *yes/no* question with rising intonation but without inversion especially under conditions that require guessing for clarifying uncertainty and confirming information.

from three occasions on which the frequency of inversion falls below 80% (70.6%, 77.8%, and 78.9% respectively), the frequency of inversion for all the other 57 conversations/speech samples range between 81-100%. Interestingly, one of the three subjects with whom the native speaker produced the lowest frequency of inversion in conversation produced the highest frequency of inversion (100%), one produced a higher frequency (86%) and the other one a lower frequency (71%). In fact, the average percentage (81.3%) of the three subjects is higher than that of all the subjects under study (76.7%). Refer to Table 4.1.5 for the comparative results.

Table 4.1.5 Token Counts of *Yes/no* Inversion for Native Speaker and Three Subjects

Participants	Total YNQ	VS in YNQ	VS in YNQ (%)
M8	2	2	100
NS (to M8)	17	12	70.6
K6	7	6	86
NS (to K6)	18	14	77.8
M10	7	5	71
NS (to M10)	19	15	78.9

In light of this evidence, we can conclude that the experimental conditions with particular reference to the effect of the native language speaker did not bias the results in favour of the universal. It is also clear that, overall, the native speaker inverts in *yes/no* questions significantly more frequently than interlanguage speakers; in contrast, interlanguage speakers vary from individual to individual. However, whatever variation there may be, interlanguage development conforms to the universal constraints of the implicational hierarchy.

An important point to be made here is that the data in this study, like those reported by Eckman et al (1989), overwhelmingly support the implicational hierarchy in question regardless of the subjects' English proficiency levels and first languages. Table 4.1.6 provides count details of Features A, B and C for the subjects in terms of these two factors.

Table 4.1.6 Token Counts and Percentages for the Three Features by Levels & L1s

Features	First Languages*								English Levels**		
	M (16)	J (7)	K (8)	V (5)	T (7)	In (3)	R (5)	B (1)	Low (16)	Mid (18)	High (18)
Total <i>wh</i> questions	43	19	28	16	21	12	18	9	44	59	63
A – <i>Wh</i> Fronting	43	19	28	16	21	12	18	9	44	59	63
B – <i>Wh</i> Inversion	39 91%	18 95%	27 96%	15 94%	19 90%	11 92%	16 89%	8 89%	42 95%	54 92%	57 90%
Total <i>yes/no</i> questions	104	44	49	25	46	21	24	1	83	136	94
C – <i>Yes/no</i> Inversion	81 78%	36 82%	38 78%	23 92%	30 65%	14 67%	18 75%	1 100%	64 77%	99 73%	78 83%

* First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish, French and Italian. The subject speaking Bangla is presented for completeness but not for analysis since it is inapplicable in terms of comparison between groups. The same applies to the subsequent tables that tabulate data for comparison between L1 groups.

** English proficiency level: low level = levels a & b; mid level = level c; high level = level d. Names of low-, mid-, and high-level groups are conveniently used for analysis, and do not necessarily indicate the proficiency difference as their respective names may suggest (refer to 3.1.3 and Appendix II for details).

From the table, it can be seen that with respect to the factor of first languages, all L1 groups demonstrate conformity to the implicational universals. It is significant that the subjects whose first languages such as Mandarin, Japanese and Korean do not systematically exhibit either the implicans or implicatum of the two implicational universals perform in the same way as those Romance-speaking subjects whose first languages do show systematically both implicantia and implicata of the universals. This shows that what is responsible for the acquisition of questions in relation to implicatia and implicata of the universals at issue is not language transfer but what Eckman et al refer to as *Interlanguage Structural Conformity Hypothesis*, which is presented in (2).

- (2) *Interlanguage Structural Conformity Hypothesis (Interlanguage SCH)*. All universals that are true for primary languages are also true for interlanguages.

Apart from the overall conformity to the universals by all L1 groups, there do exist some differences as to the degree of conformity to the implicational universal “*Yes/no* Inversion implies *Wh* Inversion”. The Vietnamese-speaking group gains a much higher percentage of Feature C – *yes/no* inversion (92%) against the other groups, well above the average percentage (77%), while Thai- and Indonesian-speaking groups both have

low percentages (65% and 67% respectively). Interestingly, the only four subjects who have less-than-five-years experience of learning English belong to the Vietnamese-speaking group in which three learned Russian as a foreign language in their secondary schooling. It is amazing that the Vietnamese group ($N = 5$) with an average length of learning English much less than that for all the subjects (4.8 versus 10.5 years), could have performed so well. That is, they achieved higher English proficiency in general (4 belong to the high-level group; a group with the highest English proficiency level on average), and better results of the task than the other groups in particular. Our suggestion for this phenomenon is that they were highly motivated and attended much more to form than the other groups as a function of the conditions under which they learned English – mainly via self-study and/or night classes in their home country. The experience of already learning a foreign language, Russian (similar to English in question formation), might also have enabled them to readily compare question structures in Russian and English, and to attend and acquire the form of English questions easily. Whatever the reason, one point is clear: all groups conform to the implicational universals no matter how high or low were the percentage scores of each group.

The fact that three groups of subjects with different English proficiency levels performed similarly in the direction of structural conformity to the universals indicate that interlanguage developmental factors do not affect the validity of the universals. In other words, the implicational universals in question hold true for interlanguage of the subjects whether they are at the intermediate, upper-intermediate or advanced levels. This can also be demonstrated by the performances of the two-round participants S1 and M7. In the first round of data collection, S1 was at the preliminary level (GE 2) and M7 at the advanced level (ETP 6); six months later, S1 was at the intermediate level (GE 4B) and M7 had just finished the advanced language training course, passed the IELTS test and received a full university offer. Their results are provided in Table 4.1.7.

**Table 4.1.7 Token Counts (for S1 and M7) for the Three Features in
“Yes/no Inversion implies *Wh* Inversion which in turn implies *Wh* Fronting”**

1	2	3	4	5	6	7	8	9
	<u>Yes/no Questions</u>			<u>Wh Questions</u>			<u>Wh Questions</u>	
	Total	VS in YNQ		Total	VS in WHQ		#wh in WHQ	
Subject	YNQ	No.	%	WHQ	No.	%	No.	%
S1a	8	4	50	3	2	67	3	100
S1b	6	5	83	4	4	100	4	100
M7a	6	5	83	3	1	33	3	100
M7b	9	8	89	1	1	100	1	100

From the table, we can see that developmental factors do seem to make some difference in the results with respect to “Yes/no Inversion implies *Wh* Inversion” with the subjects improving their performances in terms of higher percentages for both implicans and implicatum of the universal. For S1, however, his improvement does not change the pattern of implicans and implicatum; his performances in both rounds comply with the universal constraints. For M7, on the other hand, the change from her first-round disconfirming data for the universal to her second-round confirming data can not be surely attributed to a function of performance improvement because of the scanty data of implicatum (only one instance) in the second round.

As has been argued thus far, our methods are valid and sound despite a small sampling of questions from each subject, and our data comply with the implicational hierarchy stated in terms of Hypothesis 1 despite a significant exception to “Yes/no Inversion implies *Wh* Inversion”. The outcome of this study is taken to confirm the results in Eckman et al (1989) in that findings in both studies are in compliance with the universal constraints stated in the *Interlanguage Structural Complexity Hypothesis* (Eckman et al 1989: 195). Specifically, the ceiling effect of sentence-initial *wh* words/phrases on the “*Wh* Inversion implies *Wh* Fronting” universal, which Eckman et al doubt may be caused by high English proficiency levels of their subjects, is also attested in this study which involves subjects with various levels of English proficiency. Similarly, the “Yes/no Inversion implies *Wh* Inversion” universal has one exception in this study as well as in theirs.

Notwithstanding, their tentative explanation for the exception in processing terms does not seem to stand to reason. They noted that their suggested processing account runs

counter to the implicational universal, and that such processing ease should have resulted in more exceptions to the universal, rather than only one. In effect, their explanation seems to be paradoxically stated: if the universal holds, the processing mechanism wanes; if the processing account prevails, the universal fails. Whichever is the case, the validity of the implicational universal is called into question because universals are inseparably related to the ease with which the human mind processes language. In fact, universals and processing are supportive (cf. Hawkins 1994, 1999) rather than contradictory to one another. In view of this, a proposed processing account different from Eckman et al's is put forth.

Consider the four sentences in (3):

- (3) a. *What_i* [has he read *O_i*?]
b. Has he read NP[*the* novel]?
c. Has he read NP[horror *novels*]?
d. Has he read NP[19th-century horror *novels*]?

The structure of the *wh* question in 3a differs from that of the *yes/no* questions in 3b, 3c and 3d in that the former contains a "filler-gap domain" (Hawkins 1999) in which the filler *What* is matched by its co-indexed gap following *read*, whereas the latter three do not.⁴² Since they belong to different types of structures, it seems hard to compare their respective processing difficulties. However, if viewed in terms of the number of nodes to be processed in a constituent structure that is valency complete, the difference between 3a and 3b, 3c and 3d is clear. That is, the complete valency domain for 3a contains the same number of nodes as that for 3b in which *the* constructs the NP, but contains fewer nodes than 3c and 3d in which *novels* constructs the NP. Crucial to processing load, as has been suggested, might be the number of nodes to be processed in a complete valency domain rather than the number of syntactic operations undergone by 3a versus 3b, 3c and 3d (see Eckman et al 1989 for the latter point). If this is true, then *wh* questions are no more difficult than *yes/no* questions to process because, generally, the former have the same number of nodes or fewer nodes to process than the latter as shown by the examples in (3). This seems to be a feasible processing explanation for both the implicational

universals which are formulated via examination of primary languages, and the interlanguage data in both studies which conform predominantly to the implicational universals. The one exception in both studies then may be best accounted for by the internally unstable and variable nature of interlanguage systems.

4.1.5. Summary

The task of conversation in this study focuses on testing two implicational universals or one implicational hierarchy as stated in the form of *Hypothesis 1*, which is construed in terms of relative frequency rather than absolute arbitrarily-determined frequency thresholds, against spontaneous speech of interlanguage speakers. It is borne out that, similar to the findings in Eckman et al (1989), our data fully support "*Wh* Inversion implies *Wh* Fronting" without exception and predominantly support "*Yes/no* Inversion implies *Wh* Inversion" with only one exception. More significantly, this study supports *Interlanguage Structural Conformity Hypothesis* (Eckman et al 1989: 195) regardless of factors such as first language and English proficiency of L2 learners. All in all, the findings of this study are yet another piece of evidence for the claim that implicational universals that hold true for first languages should also hold true for second languages.

4.2. Results regarding the Acquisition of Relative Clauses with Reference to the Accessibility Hierarchy

This section deals with the analysis and discussion of results regarding English restrictive relative clauses with reference to the Accessibility Hierarchy in the interlanguage data of the study. Analysed and discussed are results of various RC-related non-standard forms (4.2.5) as well as results of relativisation from the tasks of elicited repetition (4.2.1.), sentence combination (4.2.2.), grammaticality judgment (4.2.3.), the conversation and subjects' written essays (4.2.4.). This is followed by a summary (4.2.6.).

⁴² Refer to section 2.1.3. for a detailed discussion about Hawkins' processing theory with particular reference to filler-gap domains.

4.2.1. Results from the Repetition Task

Recall (3.2.2.) that the task of elicited repetition was designed to test the validity of the Accessibility Hierarchy with interlanguage data in terms of *Hypothesis II*, which is restated in (4).

- (4) **Hypothesis II:** In the task of elicited repetition of English restrictive relative clauses, L2 learners' correct instances of repetition on SU position will be greater than or as many as those on DO position, which will, in turn, be greater than or as many as those on IO position, and so on for the other lower positions on the Accessibility Hierarchy, regardless of the learners' English proficiency levels and first language backgrounds.

The task items (see **Appendix III**) are made up of twelve stimulus sentences with each one containing a relative clause. All the twelve relative clauses have the same grammatical head noun – SU in the matrix clause. Each of the six relativised positions on the Accessibility Hierarchy is represented by two sentences (SU – S3 & S12; DO – S1 & S10; IO – S6 & S9; OBL – S5 & S8; GEN – S4 & S7; OCOMP – S2 & S11). The subjects' repetition was tape-recorded, transcribed, and then classified according to interlanguage features. Appendix V presents definitions for these features with illustrative examples and a table which summarises the subjects' performances in this task. Due to the fact that each of the six relativised positions are represented by only two instances in the task, it is therefore difficult to make a comparison between the six positions in terms of number of correct instances on an individual basis. As a result, the comparison is made in relation to groups by gender, first language and English proficiency level as well as to the overall results. The aggregate results of the task are presented in Table 4.2.1.

Table 4.2.1 Aggregate Counts for Different Features in the Repetition Task

Features	SU:		DO:		IO:		OBL:		GEN:		OCOMP:		Total
	S3	S12	S1	S10	S6	S9	S5	S8	S4	S7	S2	S11	
1. Correct repetition	35	48	15	11	5	3	3	4	17	12		2	155
2. Near correct repetition	5	6	5	4	5		1	2	3	6		2	39
3. Failed repetition	11	4	34	39	28	42	39	44	22	24	39	44	370
4. Incorrect RC marker				2	4	2	5	1	4	3		1	22
5. Different meaning			2	2	2	2							8
6. Omission of obligatory RC marker	8		1			1							10
7. Different relativisation			3	2	16	11	12	10	14	15	21	10	114
8. Topic-comment structure	1	2		1		1	2		2			1	10
9. Resumptive words				1					1	1			3

The table shows that, as predicted by *Hypothesis II* with the exception of GEN position, instances of *correct repetition* on SU position are more than those on DO position, which are more than those on IO and OBL positions, which in turn are more than those on OCOMP position. Collapsing feature 1 (*correct repetition*) and feature 2 (*near correct repetition*), the same pattern persists in percentage terms: SU (78%) > DO (29%) > IO/OBL (10%) > OCOMP (3%) ('>' means 'more accessible than').⁴³ The subjects' performances on GEN position, though, are in compliance with some of the findings in the literature (e.g. Gass 1979; Li and Li 1994). In Gass (1979), the L2 learners of English generally performed better on GEN position than on DO and IO/OBL positions on the Accessibility Hierarchy in a sentence combination task, and in Li and Li (1994), the Chinese learners of English treated GEN position as more accessible to relativisation than IO and OBL positions in a similar task. Overall, the results in the repetition task support the Accessibility Hierarchy in terms of number of (near) correct instances of repetition.

⁴³ Comrie (1981) uses 'non-direct object' to include both IO and OBL. Hawkins (1999) takes the same position, explaining IO and OBL positions as sharing the same filler-gap domain that contains nine nodes to be processed. Results in the repetition task of this study also show that these two positions are similar across all categories except for the strategy of using the same relativised position to express different meaning (feature 5). Though task items are provided separately for these two positions, they are treated hereinafter as belonging to the same relativised position from the processing perspective in terms of the same size in their respective filler-gap domains (cf. Hawkins 1999).

Strikingly, relativisation on SU is much more accessible than relativisation on the other positions in that it not only accounts for most correct instances of repetition, but it has the fewest instances of *failed repetition* (15) and no exemplifying instances in four types of non-standard forms (features 4, 5, 7 and 9).⁴⁴ Each of the other positions has at least three times as many instances of *failed repetition* (DO: 73; IO: 70; OBL: 83; GEN: 46; OCOMP: 83).

Even more interesting is the fact that in terms of repetitions that involved restructuring of the sentence, the lower the relativised position on the hierarchy, the more the instances of this feature were exemplified. Moreover, almost all instances of restructuring involved relativisation on higher positions than those original relativised positions in the stimulus sentences. For example, the pattern for instances of different relativisation is: SU (0) < DO (5) < IO (27) and OBL (22) < GEN (29) < OCOMP (31) ('<' means 'less than'), which indicates that complex structures are more prone to restructuring to less complex structures, rather than vice versa. This is even clearer with a close look at the direction to which the instances of different relativisation move. Table 4.2.2 provides the results of different relativisation.

Table 4.2.2 Token Counts of Different Relativisation in the Repetition Task

Different Relativisation		DO:		IO:		OBL:		GEN:		OCOMP:		Total
		S1	S10	S6	S9	S5	S8	S4	S7	S2	S11	
On a lower position: GEN						1						1
On a higher position:	a. SU	3	2	16	2	11	5	14	15	21	8	97
	b. DO				9		5				1	15
	c. GEN										1	1
Total		5		27		22		29		31		114

From the table, it is clear that an overwhelming majority of instances of different relativisation (113 out of 114 instances) fall on positions higher than those in the stimulus sentences with only one exception from S5. The exception is from OBL ('The girl who Sue wrote the story with is proud') to GEN-SU ('The girl whose friends read story is proud') – a lower position on the hierarchy, which, however, is consistent with

performances as regards correct repetition in the task. The subjects produced far more (near) correct instances of repetition on GEN position (38) than on IO/OBL positions (13/10).

In closing, the Accessibility Hierarchy is generally supported by the subjects' performances in the repetition task with the exception of GEN position. This exception may be due to the unique characteristic of the English genitive relative marker *whose*, which is elsewhere reported to be more accessible to relativisation than the positions of IO/OBL and even DO (e.g. Gass 1979; Li and Li 1994). This is discussed further in Chapter 6.

Now we examine whether the subjects' English proficiency levels and first languages may affect *Hypothesis II*. Results are provided in Table 4.2.3.

Table 4.2.3 Token Counts and Percentages for (Near) Correct Repetition by Levels & L1s

Relativised positions of (near) correct repetition*	English Levels			First Languages**							
	Low (20)	Mid (20)	High (20)	M (18)	J (10)	K (9)	V (7)	T (7)	In (3)	R (5)	B (1)
Total number	45 19%	54 23%	95 40%	68 32%	26 22%	21 19%	21 25%	27 32%	11 30%	13 22%	7
a. SU	27	30	37	30	16	11	11	13	5	6	2
b. DO	9	8	18	11	5	3	5	5	2	2	2
c. IO/OBL	0/0	2/4	11/6	5/5	0/0	1/2	1/0	2/1	2/0	1/2	1/0
d. GEN	8	10	20	15	4	4	4	5	2	2	2
e. OCOMP	1	0	3	2	1	0	0	1	0	0	0

* Features of correct repetition and near correct repetition are collapsed for analysis.

** First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish, French and Italian. The figures within parentheses indicate the number of subjects in each group.

Results show that in terms of (near) correct repetition, the relation between different level groups and their respective performances is significant, i.e. the high-level group scored significantly better than both the low-level group and mid-level group ($\chi^2 = 30.07$, $df = 2$, $p < .001$).⁴⁵ Nonetheless, this does not alter the predicted order of (near) correct

⁴⁴ Detailed analysis of RC-related non-standard forms in different tasks is provided in 4.2.5.

⁴⁵ The statistical procedure of Chi-square (see Hatch and Lazaraton 1991: 393-424) is used for some of the nominal data of the study. Chi-square for two-way designs with the significance level of .05 is applied to this and subsequent relevant sets of data to test whether the relation between two variables is significant.

repetition on the relativised positions for each group. Although there seems to be an exception in which OCOMP has one correct instance with IO/OBL having none for the low-level group, this exception can be disregarded for lack of disconfirming evidence.⁴⁶ In fact, as compared with the high-level group and even the mid-level group, the low-level group found it extremely hard to correctly repeat sentences that contain relativisation on IO/OBL and OCOMP positions. With respect to first language, although different language groups vary in their absolute scores, all L1 groups follow the predicted ordering pattern of relativisation in terms of instances of (near) correct repetition on each relativised position.

To sum up, *Hypothesis II* is supported by the results from the repetition task of the study, or in other words, the Accessibility Hierarchy is proved to be valid in predicting L2 performances regarding comprehension in the repetition task regardless of L2 learners' English proficiency levels and first languages.

4.2.2. Results from the Sentence Combination Task

The task of sentence combination (see 3.2.3.) was also designed to test the Accessibility Hierarchy but in L2 learners' interlanguage production in the form of *Hypothesis III*, which is re-stated in (5).

- (5) **Hypothesis III:** In the task of sentence combination involving English restrictive relative clauses, L2 learners' correct instances of combining sentences on SU position will be greater than or as many as those on DO position, which will, in turn, be greater than or as many as those on IO position, and so on for the other lower positions on the Accessibility Hierarchy, regardless of the learners' English proficiency levels and first language backgrounds.

⁴⁶ As mentioned before, the occurrence of one instance does not systematically confirm or disconfirm a pattern. Moreover, the correct instance on OCOMP position for the lower-level group is only one of near correct repetition. So is the same explanation made for the Japanese-speaking group under the same circumstances.

The task items are made up of fourteen pairs of sentences and the subjects are required to combine each pair into a single sentence with a relative clause following a sample example. Each of the six relativised positions on the Accessibility Hierarchy is represented in two instances (SU – S1 & S14; DO – S2 & S13; IO – S3 & S12; OBL – S4 & S11; OCOMP – S7 & S9) except GEN, which is further divided into GEN-SU (S5 & S6) and GEN-DO (S8 & S10). The two instances representing the same relativised position have different grammatical head nouns in the matrix clause – in SU and DO positions respectively. The subjects' performances in the task were sorted out and classified according to different interlanguage features. **Appendix VI** presents definitions for these features with illustrative examples and a table which summarises the subjects' performances in this task. Like the repetition task, each of the six relativised positions is also represented by only two instances in the task, it is therefore difficult to make a comparison between the six positions in terms of number of correct instances on an individual basis. As a result, the comparison is made in relation to groups by the subjects' English proficiency levels and first languages as well as to the overall results. The aggregate results of the task are presented in Table 4.2.4.

Table 4.2.4 Aggregate Counts for Different Features in the Sentence Combination Task

FEATURES	SU		DO		IO		OBL		GEN-SU		GEN-DO		OCOMP		TOTAL
	1*	14	2	13	3	12	4	11	5	6	8	10	7	9	
1. Predicted RC embedding	55	41	25	18	12	11	36	12	38	31	15	21	14	10	339
2a. Relativization on a similar position		3	8												11
2b. Relativization on a higher position			18	33	30	34	17	36	6	5	27	26	33	36	301
3. No adjacency to the head noun		5			1	1				5	1	1	2	3	19
4. Inaccessible head noun									2		4				6
5. Incorrect relative markers		5	2	5	6	2		3	2	7		2	1	2	37
6. Topic-comment structure							1	1		1	1	1	1	1	7
7. Use of resumptive words				2	1	1	1	1	2	7			2	4	21
8. Omission of obligatory RC marker		1		1		1		4					1		8
9. Other errors concerning RCs			1		1									1	3
10. Failure to supply a RC	1	2	4	2	9	6	2	1	10	8	13	9	4	5	76

* Figures in this row indicate individual task item numbers.

Like the repetition task, the aggregate results of the combination task generally support *Hypothesis III* with the exception of GEN position. That is, as regards the feature of *predicted RC embedding*, the average percentage of instances on SU position (86%) is higher than that on DO position (38%), which is higher than that on IO/OBL positions (32%), which is in turn higher than that on OCOMP position (21%). The order of accessibility to the relativised positions in terms of *predicted RC embedding* can then be presented as: SU (86%) > GEN (47%) > DO (38%) > IO/OBL (32%) > OCOMP (21%) ('>' means 'more accessible than'). The subjects' performances on GEN position are, again, similar to the results of the repetition task: better than not only those on OCOMP but those on DO and IO/OBL positions as well (cf. Gass 1979; Li and Li 1994). Overall, the results in the sentence combination task support the Accessibility Hierarchy in terms of predicted relative clause embedding.

Relativisation on SU proves to be much more accessible than the other positions. Not only instances of *predicted RC embedding* on this position are of the highest percentage (86%), the preference for this position is also shown in the subjects' performances in relativising on GEN position. The instances of *predicted RC embedding* on GEN-SU position (SU in the relative clause) are more than those on GEN-DO position (DO in the relative clause): 69 versus 36 (or 62% versus 32%). Furthermore, 11 out of the 36 instances of *predicted RC embedding* on GEN-DO position, are actually shifted onto GEN-SU position via passive voice as shown in (6).

(6) a. **Predicted RC embedding for task item 8:**

The patient whose disease the doctor regarded as incurable committed suicide.

Actual example:

The patient whose disease was regarded as incurable by the doctor committed suicide.

b. **Predicted RC embedding for task item 10:**

My son saw the lady whose house we bought last week.

Actual example:

My son saw the lady whose house was bought by us last week.

This can be well accounted for in terms of processing load the two relativised positions – GEN-SU and GEN-DO – carry. The minimal filler-gap domain for the former contains

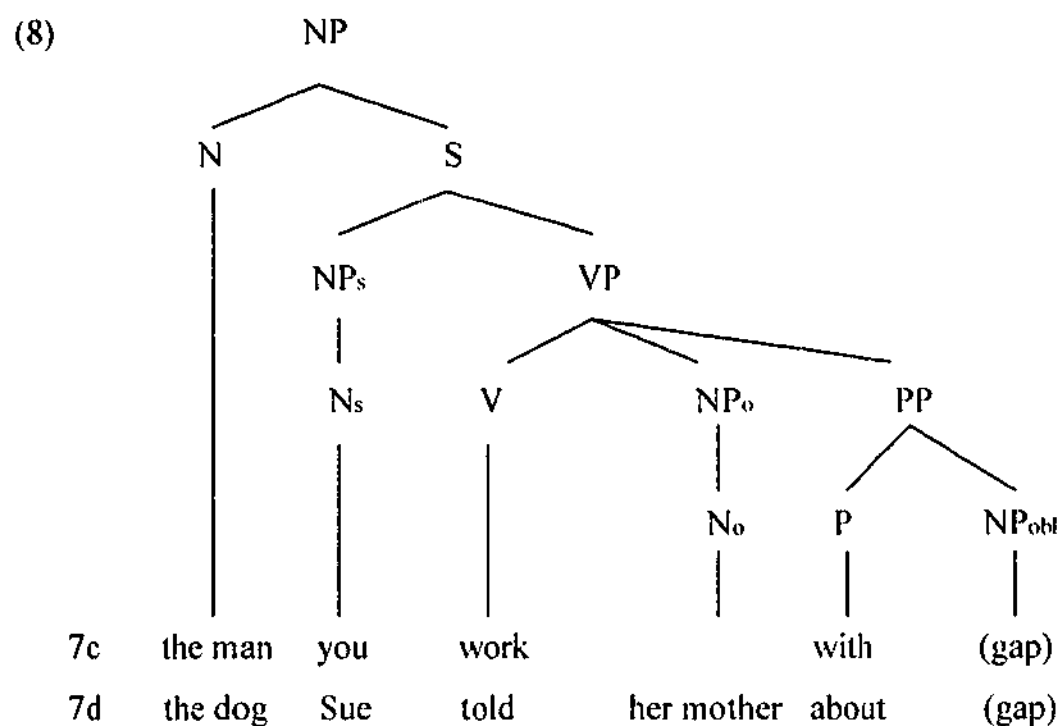
nine nodes to be processed, while that for the latter contains eleven nodes (Hawkins 1999: 255). To facilitate processing, the human processor prefers the filler-gap domain to be as small as possible. Therefore, whenever a structural change occurs, the change moves from more complex to less complex structure, rather than vice versa. This is borne out by the subjects' performances on GEN positions: no change occurs among the instances of predicted relativisation on GEN-SU, and among those on GEN-DO, changes do occur in favour of processing preference, which is achieved by means of passive voice without changing the basic meaning.

Regarding results from the table, there are a number of interesting observations. First, with respect to the two task items intended for each relativised position, the subjects tended to perform better on the one that does not have "interruption" (matrix NPs and VP are separated by a relativising clause) (Sadighi 1994: 145) than the one that does with only one exception. That is, when scores differ between two task items for the same relativised position, the difference tends to favour the "non-interrupted" one (SU – 55: 41; DO – 25: 18; IO – 12: 11; OBL – 36: 12; GEN-SU – 38: 31; GEN-DO – 21: 15) except for OCOMP position (10: 14). Therefore, the "interruption" account (Sadighi 1994) generally holds for the data with one exception. However, this account alone can not adequately address the following question regarding IO/OBL positions: why is there only a single token difference for IO position but a triple-percentage difference for OBL position, both of which occur under similar "interruption" circumstances? This seems puzzling and a possible plausible answer is provided in the next point.

Second, although the subjects' overall performances in relativising on IO/OBL positions are predictably in compliance with *Hypothesis III*, their performances in task item 4 (36; 64%) did differ greatly from those in task items 3 (12; 21%), 11 (12; 21%), and 12 (11; 20%). Look at the predicted sentence combination for these task items in (7).

- (7) a. Item 3: The girl saw *the actor*_i [*who(m) she sent a letter to* O_i].
 b. Item 12: *The boy*_i [*who(m) the girl lent her car to* O_i] broke his words.
 c. Item 4: The detective questioned *the man*_i [*who(m) you work with* O_i].
 d. Item 11: *The dog*_i [*which Sue told her mother about* O_i] won a prize.

The subjects' performances on IO position 7a and 7b and OBL position 7c and 7d can be explained in processing terms. The filler-gap domain for 7a, 7b and 7d contains, in fact, eleven nodes to be processed, whereas that of 7c contains nine nodes to be processed. For example, the configurational assumptions of the filler-gap domains of 7c and 7d are illustrated in (8).⁴⁷



The filler-gap domain of 7c contains nine nodes (N, NP, N_s, NP_s, S, V, VP, P and PP), while that of 7d has two more nodes (N_o and NP_o) to be processed. It seems that regarding relativisation on IO/OBL positions, it does not matter so much for the facilitation of processing whether the filler (head noun) is a matrix subject or matrix object. What really matters in this regard is the size of a filler-gap domain in which a relative clause is processed. It is, therefore, not surprising that 7c elicited far more instances of the predicted RC embedding than 7a, 7b and 7d.

⁴⁷ The illustrative tree representation "does not assume any additional WH-movement within the relative clause (i.e. no relative pronouns, simply a complementizer or other subordination indicator which has been omitted here since it is constant across all the positions)" (Hawkins 1999: 254), hence the filler copies *who(m)* and *which* omitted. NP_s = subject, NP_o = direct object, and NP_{obl} = oblique object.

Finally, the most common type of non-standard forms, again, is the feature of *different relativisation* – relativisation on a position different from the predicted RC embedding (312 instances in total). Among the instances of this category, eleven (SU: 3 and DO: 8) relativise on a similar position on the hierarchy, which indicates that relativisation variation to similar positions is confined to more accessible SU and DO positions on the accessibility hierarchy. Moreover, there are no instances of relativising on a position lower than the one predicted on the hierarchy, and all the other instances of *different relativisation* (301) involve relativising on a position higher than the one predicted on the hierarchy. Table 4.2.5 provides results of relativisation on higher positions in the sentence combination task.

Table 4.2.5 Token Counts of Relativisation on Higher Positions in the Sentence Combination Task

Relativisation on higher positions	DO		IO		OBL		GEN-SU		GEN-DO		OCOMP		Total
	S2	S13	S3	S12	S4	S11	S5	S6	S8	S10	S7	S9	
SU	18	33	19	34	12	36	6	5	11	8	33	18	233
DO			11		5					2		8	26
IO									16	13			29
OBL										3			3
GEN-SU												10	10

Obviously, SU position is, again, the most accessible one to relativisation with a total of 233 instances across all the task items in the table. Relativisation on DO position (26 instances) occurs across 4 task items (S3, S4, S9 and S10), on IO position (29 instances) across 2 task items (S8 and S10), and relativisation on both OBL (3 instances) and GEN-SU position (10 instances) occurs in only one task item (S10 and S9 respectively). Even though instances of different relativisation on positions lower on the hierarchy such as OBL and GEN-SU do occur, these positions are still higher than the predicted relativised positions, i.e. GEN-DO and OCOMP positions respectively. Interestingly, passive voice, which is the sole means used by the subjects in instances of shifting from GEN-DO position to GEN-SU position, is, again, employed mainly in instances of shifting relativisation from DO to SU positions on the hierarchy. For example, among all instances of relativisation to the highest position (i.e. SU) on the hierarchy via passive voice (35 in all), twenty-three instances come from DO position with only 2 from IO, 2

from OBL, 2 from GEN-SU and 6 from GEN-DO. This demonstrates that passive voice plays an important role in shifting relativised positions from DO to SU. Importantly, all the evidence concerning different relativisation in the sentence combination task shows that it lends strongly to the validity of the Accessibility Hierarchy in terms of direction of relativisation variation.

Now we turn to examining the subjects' performances on the sentence combination task with respect to English proficiency levels and first languages, the results of which are provided in Table 4.2.6.

Table 4.2.6 Token Counts and Percentages for *Predicted RC Embedding* by Levels & L1s

Relativised positions of <i>Predicted RC</i> <i>Embedding</i>	English Levels			First Languages*							
	Low (20)	Mid (20)	High (16)	M (18)	J (10)	K (7)	V (6)	T (7)	In (3)	R (4)	B (1)
Total number	119 43%	131 47%	89 40%	113 45%	78 56%	53 54%	41 49%	32 33%	11 26%	7 13%	4
a. SU	32	36	28	27	18	12	12	14	6	5	2
b. DO	15	18	10	11	11	9	4	6	1	1	0
c. IO/ OBL	8 17	9 19	6 12	7 16	9 11	5 8	1 6	1 4	0 1	0 1	0 1
d. GEN-SU	27	23	19	28	16	9	9	5	2	0	0
GEN-DO	11	18	7	16	8	6	5	0	0	0	1
e. OCOMP	9	8	7	8	5	4	4	2	1	0	0

* First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish, French and Italian. The figures within parentheses indicate the number of subjects in each group.

Different proficiency level groups also varied in terms of their respective absolute scores. Unlike the repetition task, it is the mid-level group that achieved the best score (131; 47%) in instances of *predicted RC embedding* with the low-level group (119; 43%) scoring, surprisingly, a bit better than the high-level group (89; 40%). The high-level group, however, has less within-group variation than the other two groups. For example, the numbers for the subjects achieving 10 or more instances of *predicted RC embedding* within groups of high level, mid level and low level are 1, 5 and 5 respectively, and those achieving 2 or less are 1, 3 and 5 respectively. Notwithstanding, it is surprising that there are more high-scorers in this task in the low- and mid-level groups than in the high-level group. Examined in their performances on the repetition task, these high-scorers are found to be ranked as average scorers or higher-than-average-but-lower-than-top scorers

in terms of (*near*) *correct repetition*. In addition, five of the ten high-scorers in the low- and mid-level groups belong to the Japanese-speaking group, which is mainly made up of the Japanese university students studying English here as part of their course requirements. All this may indicate that these high-scorers with many coming from the Japanese-speaking group may be better at producing written forms rather than at comprehending spoken utterances. Despite all these differences, it is, again, significant that all groups preformed in the predicted way – in conformity with the Accessibility Hierarchy, quite expectedly, with the exception of GEN position.

In view of the subjects' performances regarding *predicted RC embedding* by different L1 groups, all L1 groups generally support *Hypothesis III*, yet again with exception of relativisation on GEN position. The results of relativisation on GEN position in this task, similar to the results of relativisation on this position in the repetition task, are again consistent with the findings reported in Gass (1979) and Li and Li (1994) (see 2.2.3.2 and 4.2.1 for details). However, in the Vietnamese-speaking group, the instances of relativisation on OCOMP position (4) are roughly the same as those on DO position (4) and IO/OBL positions (3.5); in the Mandarin-speaking group, those on IO/OBL positions (11.5) are marginally better than those on DO (11). Although these cases do not disconfirm *Hypothesis III*, why these two groups performed differently from other groups in these respects remains to be explained. Now we examine the Vietnamese-speaking group in detail to see why this might be the case. Results of the task for this group are provided in Table 4.2.7.

Table 4.2.7 Vietnamese Subjects' Performances in the Sentence Combination Task

No	Code name	SU		DO		IO		OBL		GEN-SU		GEN-DO		OCOMP	
		S1	S14	S2	S13	S3	S12	S4	S11	S5	S6	S8	S10	S7	S9
1	V1a	1	1	2a	10	2b	10	1	2b	1	3	10	2b	2b	2b
2	V2	1	1	2b	2b	2b	2b	2b	2b	1	2b	1	1*	2b	2b
3	V3	1	1	1	1	2b	2b	1	1	1	1	1*	1*	1	1
4	V4	1	1	1	2b	2b	2b	2b	2b	1sp	2b,6	2b	2b	2b	1
5	V5	1	1	2a	2b	2b	2b	1	2b	1	1	2b	2b	2b	2b
6	V6	1	1	2a	1	2b	1	1	1	1	1	2b	1	2b	1

From the table, it is clear that performance on the task item concerning DO position (S2) has three instances of *relativisation on a similar position* (feature 2a) different from the predicted sentence combination, which are all supplied in (9).

(9) **Task item 2 (S2):** The dog ate the bone. The boy found the bone.

Predicted RC embedding: The dog ate the bone that/which the boy found.

Three instances of feature 2a:

- a. The boy found the bone that the dog ate.
- b. The boy found the bone which the dog ate.
- c. The boy found the bone which the dog ate.

Deviating from the predicted RC embedding, the three instances, though, match it perfectly in terms of the relativised position – DO on the hierarchy. In this sense, the three instances can be ranked as equal to those of *predicted RC embedding*, because the three Vietnamese-speaking subjects know exactly how to relativise on DO position though they reversed the order of the clauses and the resulting relative clauses were slightly different in meaning. This may suggest that these Vietnamese subjects may have attended more to form than to meaning in dealing with the task item. Viewed in this way, the DO position for the Vietnamese-speaking group, in fact, comprises more instances of *predicted RC embedding* (7 in all) than those for IO/OBL and OCOMP positions (3.5 and 4 respectively), hence more accessible to relativisation. Similarly, the Mandarin-speaking group also has 3 instances of feature 2a (see **Appendix VI**), and therefore also has more instances of *predicted RC embedding* on DO position (14 in all) than on IO/OBL positions (11.5). Furthermore, within the Vietnamese-speaking group, the subject, V3, treated the two instances on OCOMP position on a par with those on OBL position, since *whom* instead of *who* was used in both instances of relativisation on OCOMP position. This may help to explain similar performances pertaining to relativisation on OCOMP position and on IO/OBL positions for the group. It appears now that both the Vietnamese- and Mandarin-speaking groups performed similarly to the other groups, all conforming to the way predicted by the Accessibility Hierarchy in the sentence combination task.

In closing, *Hypothesis III* is overwhelmingly supported by the data in the sentence combination task. In spite of some variations thus far examined, they do not seem to contradict the hypothesis. We can conclude, on the basis of our analysis, that the Accessibility Hierarchy is generally valid in predicting L2 performances regarding RC production in the sentence combination task regardless of L2 learners' English proficiency levels and first languages.

4.2.3. Results from the Grammaticality Judgment Task

The grammaticality judgment task comprises 36 task items (stimulus sentences), five of which are designed to test subjects' comprehension of relative clauses with reference to different relativised positions. Sentences 1, 2, 5, 6 and 19 are correct sentences with each one containing a relative clause on one of the four relativised positions (SU: S1; DO: S6 & S19 with the RC marker omitted; IO: S5; GEN: S2). See **Appendix III** for reference to these sentences in the task and results of individual the subjects' performances on this task are provided in **Appendix VII**. The aggregate results of the task are presented in Table 4.2.8.

Table 4.2.8 Counts and Percentages involving RCs in the Grammaticality Judgment Task

Judgment choices (56 subjects)	Relativised Positions				
	SU	DO		IO	GEN
	S1	S6	S19	S5	S2
Acceptable	38 68%	32 57%	40 71%	26 46%	45 82%
Unacceptable	17 30%	20 36%	15 27%	28 50%	8 14.5%
Not sure	1 2%	4 7%	1 2%	2 4%	2 3.6%

Table 4.2.8 shows that the subjects' performances on relativised positions are generally in compliance with the Accessibility Hierarchy except, presumably, those on GEN position. In terms of percentages of overall correct responses to the relativised positions, the subjects' performances can be summarised as: GEN-SU (82%) > SU (68%) > DO (64%) > IO (47%) ('>' means 'more acceptable'). As compared with the repetition and sentence combination tasks, GEN position is judged to be even more acceptable than SU position in this task, and SU position with a relatively low percentage is only marginally more

acceptable than DO position (DO in S19 is marginally more acceptable than SU position). The relative percentage decline on SU position in this task may be partly due to the choice of words 'which' and 'poodle' in S1 (The dog which is running is a poodle). 'That' instead of 'which' is a better word for the *restrictive* relative clause because the former carries an undertone of 'definiteness'; besides, a few subjects asked me about the meaning of 'poodle' after the completion of the task. It seems that the low acceptability of S1 was related more to the choice of relative pronoun, and poor comprehension of the lexical item 'poodle'. Had 'that' and, say, 'mine' been substituted for 'which' and 'poodle' respectively, the percentage of the subjects' correct responses on SU position might have been increased.

However, a more principled explanation may be offered if we take a careful look at the three task items regarding relativisation on SU and DO positions (S1, S6 & S19), which are presented in (10).

- (10) S1: The dog which is running is a poodle.
 S6: The man hit the boy who you saw.
 S19: The book she bought last week is missing.

While in the repetition and sentence combination tasks, all the instances of relativisation on SU position involve a configuration of 'SU *vs* DO', the subject relativised on in this task (S1) has a different configuration, 'SU *vs*'. In terms of argument coding, the transitive subject A (agent) is coded the same (nominative) as the intransitive subject S while the transitive object P (patient) is coded differently (accusative) in nominative-accusative languages; but in ergative-absolutive languages, S and P are coded the same (absolutive) with A being coded differently (ergative) (Comrie 1978). Based on her study, Fox (1987) challenges the "subject primacy" of the Accessibility Hierarchy and proposes instead the *Absolutive Hypothesis*, which holds that "every language which has a strategy for relativizing must be able to relativize on at least S and P" (Fox 1987: 864). In the case of the repetition and sentence combination tasks, all those instances of SU relativisation lend support to the Accessibility Hierarchy rather than the Absolutive Hypothesis. In this task, however, since there are no instances of relativisation on A,

instances of relativisation on S (I1: S1) and on P (I1: S6 & S19) seem to be better accounted for by the Absolute Hypothesis, though the Accessibility Hierarchy also holds (albeit weakly). To test more thoroughly and finely which assumption holds for what kind of data it is advisable that future research in this regard should include a variety of tasks designed for testing relativisation on all these argument positions S, A and P.

Now we examine whether factors of English proficiency and L1 may affect the subjects' judgment on these items; the results of examination of these factors are provided in Table 4.2.9 and Table 4.2.10 respectively.

Table 4.2.9 Counts and Percentages of 'Acceptable' Responses involving RCs in the Grammaticality Judgment Task by Levels

Relativised Positions	English Proficiency Levels					
	Low (20 subjects)		Mid (20 subjects)		High (16 subjects)	
SU (S1)	13	65%	11	55%	14	88%
DO (S6 & S19)	24	60%	26	65%	22	69%
IO (S5)	10	50%	9	45%	7	44%
GEN (S2)	18	90%	14	74%	13	81%
Total	65	65%	60	60%	56	70%

Results from Table 4.2.9 show that the performances of different level groups on relativised positions are generally consistent with the overall results except for the mid-level group's performances concerning relativisation on SU and DO positions. That is, the mid-level group regarded relativisation on DO as marginally more acceptable than relativisation on SU. Apart from this, the overall results for the three groups are similar even if the high-level group gained a marginally higher percentage than the other two groups in regarding as acceptable the five relative clauses in the grammaticality judgment task.

Table 4.2.10 Counts and Percentages of 'Acceptable' Responses
involving RCs in the Grammaticality Judgment Task by L1s

Relativised Positions	First Languages*							
	M (18)	J (10)	K (7)	V (6)	T (7)	In (3)	R (4)	B (1)
SU (S1)	11 61%	8 80%	6 86%	4 67%	5 71%	2 67%	2 50%	
DO (S6 & S19)	30 83%	14 70%	10 71%	4 33%	7 50%	2 33%	3 38%	2
IO (S5)	5 28%	5 50%	3 43%	5 83%	5 71%	2 67%	1 25%	
GEN (S2)	15 83%	8 80%	6 86%	6 100%	5 71%	1 33%	3 75%	1
Total	61 68%	35 70%	25 71%	19 63%	22 63%	7 47%	9 45%	3 60%

* First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish, French and Italian. The figures within parentheses indicate the number of subjects in each group.

There is more variation regarding performances on relativised positions by different language-speaking groups. As expected, the Absolutive Hypothesis fares a bit better than the Accessibility Hierarchy in explaining performances by different groups, there are, however, still three exceptions: (pertaining to relativisation on DO and IO positions) for which neither of the two approaches can adequately account. IO position is lower than DO position on the Accessibility Hierarchy, and is out of the domain of 'absolutive primacy' (containing DO) according to the Absolutive Hypothesis, hence less accessible in either case. However, the Vietnamese-, Thai-, and Indonesian-speaking groups all scored higher on IO position (83%, 71% and 67% respectively) than on DO position (33%, 50% and 33% respectively).

These exceptions may well be due to interlanguage variability, the nature of the task, lack of sufficient data, or a function of the three. As compared with first languages of the subjects, their English interlanguage is in a process of less stable, more variable and ever-changing development. The exceptions at issue might be such a case in point. It may as well be that the L2 learners in the study diverge more widely in the grammaticality judgment task than other tasks owing to the extent to which they are familiar with the task and to which their knowledge of grammar varies. Insufficient data with respect to fewer items in the task for these three groups may contribute to the exceptions. Their performances on relativised DO and IO positions vary by only a few instances: the Vietnamese-speaking group by 3 instances (4 in two items for DO versus 5 in one item

for IO), the Thai-speaking group by 1.5 (7 versus 5), and the Indonesian-speaking group by 1 (2 versus 2). In the examination of the subjects' performances on relativised positions in the repetition and sentence combination tasks and even in this task on the basis of level where relatively more data are gained, no such kind of exceptions are found. This indicates that other things being equal, the more the data are gained for analysis, the more reliable and valid the results will be. Owing to the constraints of designing the task, this was, unfortunately, not achieved. With more items concerning relativisation on DO and IO positions included in this task, the results would have probably been different. Finally, all the three proposed reasons may help to explain part of the picture of the exceptions.

Interestingly, with regard to the issue of relativisation, whilst the other groups (except the Indonesian-speaking group) did similarly well in relativised positions, the Romance-speaking group performed consistently under par in this task as well as in the repetition and sentence combination tasks.

4.2.4. Results from the Conversation and Subjects' Written Essays

During the whole procedure of data collection, an English interlanguage database of spontaneous speech production in the conversation task (totaling to about 47,552 words) from 60 subjects, and of original writing production in the collected essays (totaling to 10,575 words) from 45 subjects, is established. The subjects' speech and written essays are examined and relative clauses therein are collected, tabulated, and presented in **Appendix VIII**. It is found that there are only 27 relative clauses used in the subjects' speech production with an average of one relative clause appearing in about every 1,761 words whereas there are 99 relative clauses used in their writing with an average of one relative clause in every 107 words. This indicates that the subjects rarely use relative clauses in their speech, but they do use relative clauses far more frequently in their writing.

In both their speech and writing, the subjects use relative clauses that relativise mainly on SU and DO positions with only five instances of relativisation on OBL position and one

instance on GEN position. The order of accessibility to relativisation in the subjects' speech and writing can then be presented respectively as: SU (13) > DO (10) > OBL (3) > GEN (1) and SU (79) > DO (18) > OBL (2) ('>' means 'more accessible than'). SU position, once more, is the most accessible one to relativisation with DO position having relatively high rate of accessibility especially in the speech and OBL/GEN positions having a very low rate. Viewed in terms of ergative-absolutive differentiation, the Absolutive Hypothesis also holds for the data. In fact, it accounts better than the Accessibility Hierarchy for relativisation in the subjects' speech – Absolutive (S & P: 17) > Ergative (A: 5), but not as well for relativisation in their writing – Absolutive (56) > Ergative (41). The evidence in speech lends support to the analysis of English speech on the basis of which Fox (1987) postulates the Absolutive Hypothesis. In this regard, more data is needed for testing further whether this is the case as well as whether interlanguage writing differs greatly.

In sharp contrast to the results from the repetition, sentence combination and grammaticality judgment tasks, though, relativised GEN position, on which the subjects performed extremely well in those tasks, was only exemplified with a single instance of use in the subjects' naturalistic oral and written interlanguage production. Moreover, even the only instance of relativisation on GEN position in the speech (i.e. 'Once I took listening practice, yeah, the, the tape, the woman in the tape *who* accent is terrible') involved *incorrect RC marker*. This seems to confirm the fact that relativisation on GEN (like relativisation on OCOMP) position occurs far more infrequently than relativisation on the other positions on the Accessibility Hierarchy in L2 data. It also suggests that although relativised GEN position has a cognizant prominence both in the grammaticality judgment task (involving comprehension) and in the repetition and sentence combination tasks (involving more controlled production than the subjects' free speech and writing), it may still be too complex a structure for L2 learners to use in their naturalistic L2 production. Nevertheless, both the Accessibility Hierarchy and the Absolutive Hypothesis can account for the overall results gained from the naturalistic interlanguage data of the subjects.

4.2.5. Results of Various RC-Related Non-Standard Forms

There are various relative-clause-related non-standard forms in the subjects' performances across different tasks. In the repetition and sentence combination tasks, any instances that deviate from the norm – *(near) correct repetition* or *predicted RC embedding* respectively – are regarded as non-standard forms in this study no matter whether these deviations are grammatical or ungrammatical in English. Refer to **Appendices V and VI** for the classification of the norm and types of non-standard forms in these two tasks. However, the types of non-standard forms vary with respect to grammaticality or intelligibility. By 'grammaticality', we mean that some non-standard forms such as those from *different relativisation*, *different meaning* and *topic-comment structure* are grammatical, while other non-standard forms such as those from *failed repetition*, *incorrect RC marker* and *omission of obligatory RC marker* are ungrammatical. 'Intelligibility' applies to those ungrammatical non-standard forms, amongst which some are more intelligible such as instances of *incorrect RC marker*, *resumptive words* and *no adjacency to the head noun*, and others are less so such as instances of incomplete and nonsense chunks from *failure to produce the required RC*.

In the grammaticality judgment task, on the other hand, non-standard forms refer to instances of judging as acceptable those task items that contain an RC-related non-standard form such as omission of an obligatory RC marker or a prenominal RC (see **Appendix VII**). Most deviated instances in the repetition and sentence combination tasks contain one type of non-standard form, some may contain two types, and only one instance contains three types. All these non-standard forms together with those from the subjects' speech and written essays are analysed and discussed in relation to both task variety and factors of English proficiency levels and L1 backgrounds of the subjects.

4.2.5.1. RC-Related Non-Standard Forms in Different Tasks

Results of RC-related non-standard forms are first analysed in each of the tasks in this study, and then compared across these tasks. Table 4.2.11 presents the results of different types of non-standard forms in the repetition task.

Table 4.2.11 Token Counts of Types of Non-Standard Forms in the Repetition Task

Features (non-standard forms)	SU:		DO:		IO:		OBL:		GEN:		OCOMP:		Total (537)
	S3	S12	S1	S10	S6	S9	S5	S8	S4	S7	S2	S11	
3. Failed repetition	11	4	34	39	28	42	39	44	22	24	39	44	370 69%
4. Incorrect RC marker				2	4	2	5	1	4	3		1	22 4.1%
5. Different meaning			2	2	2	2							8 1.5%
6. Omission of obligatory RC marker	8		1			1							10 1.9%
7. Different relativisation			3	2	16	11	12	10	14	15	21	10	114 21%
8. Topic-comment structure	1	2		1		1	2		2			1	10 1.9%
9. Resumptive words				1					1	1			3 0.6%

Results show that the majority of non-standard forms in the repetition task come from *failed repetition* (69%), by which subjects did not try, tried incompletely or uttered thoroughly unintelligible chunks. Specifically, on average, each subject failed six of the possible twelve instances of repetition, indicating repeating English sentences with a relative clause is a difficult task for L2 learners. An interesting fact about *different meaning* (feature 5) and *omission of obligatory RC marker* (feature 6) is that all instances of the two types of non-standard forms (8 and 10 tokens respectively) fall on higher positions (SU, DO and IO) of the hierarchy. This phenomenon may be explained in processing terms. Due to less processing load for positions higher on the hierarchy, structures of these positions are more easily perceived and comprehended than those positions lower on the hierarchy. When the subjects were unable to remember and repeat the words heard, they tended to retain structures with less structural complexity, i.e. positions higher rather than lower on the hierarchy, resulting in the same structures with choice of other words (feature 5). On the other hand, when focus is shifted onto meaning rather than form, less complex structures of relativisation (such as the highest position on the hierarchy – SU) are more likely to drop the relative markers without misinterpretation of the intended meaning, hence the case in feature 6. Omission of an obligatory relative marker for relativised positions low on the hierarchy (such as GEN) may make the

already complex structures even harder to comprehend and repeat, and may therefore result in *failed repetition*. The instances of *topic-comment structure* spread across all relativised positions; yet only one subject (K8) consistently used this structure (5 out of 10 instances of this category belong to him), which is quite idiosyncratic of his speaking style. The use of resumptive words is found only on three occasions.

Now we examine different types of non-standard forms in the sentence combination task, the results of which are provided in Table 4.2.12.

Table 4.2.12 Token Counts of Types of Non-Standard Forms in Sentence Combination Task

Features (non-standard forms)	SU		DO		IO		OBL		GEN-SU		GEN-DO		OCOMP		TOTAL (489)
	1*	14	2	13	3	12	4	11	5	6	8	10	7	9	
2. Different relativisation		3	26	33	30	34	17	36	6	5	27	26	33	36	312 64%
3. No adjacency to the head noun		5			1	1				5	1	1	2	3	19 3.9%
4. Inaccessible head noun									2		4				6 1.2%
5. Incorrect relative markers		5	2	5	6	2		3	2	7		2	1	2	37 7.6%
6. Topic-comment structure							1	1		1	1	1	1	1	7 1.4%
7. Use of resumptive words				2	1	1	1	1	2	7			2	4	21 4.3%
8. Omission of obligatory RC marker		1		1		1		4					1		8 1.6%
9. Other RC-related non-standard forms			1		1									1	3 0.6%
10. Failure to supply an RC	1	2	4	2	9	6	2	1	10	8	13	9	4	5	76 16%

* Figures in this row indicate individual task item numbers.

** Features 2a and 2b (see Table 4.2.4) are collapsed into a single one.

Unlike the repetition task, *different relativisation* in this task is the most common type of non-standard forms with 64% of the non-standard forms in this task falling in this category, a detailed analysis of which has been dealt with in the previous section (4.2.2.). *Failure to supply an RC* is the second most common type of non-standard forms (16%), which occurs in all the task items. Although less common than the above two, three types of non-standard forms are fairly common: feature 5 (*incorrect relative markers*; 7.6%) covers all relativised positions, and features 3 (*no adjacency to the head noun*; 3.9%)⁴⁸

⁴⁸ English restrictive relative clauses are usually adjacent to the head noun (non-interruption) unless they are very heavy. Although postposed relative clauses (non-adjacent to the head noun) are acceptable in spoken English, it is still not a standard practice in written English. For example, the two combined

and feature 7 (*use of resumptive words*; 4.3%) spread across most relativised positions. The least common types of non-standard forms – features 4 (*inaccessible head noun*; 1.2%), feature 6 (*topic-comment structure*; 1.4%), feature 8 (*omission of obligatory RC marker*; 1.6%) and feature 9 (*other RC-related non-standard forms*; 0.6%) are characteristic of only a couple of individuals. Unlike the repetition task in which instances of *failed repetition* are nearly all those of non-, partial or unintelligible repetition, the feature of *failure to supply an RC* in this task involves more instances of varied attempts, the results of which are provided in Table 4.2.13.

Table 4.2.13 Token Counts of *Failure to Supply an RC* in the Sentence Combination Task

Types of <i>Failure to Supply an RC</i>	SU		DO		IO		OBL		GEN-SU		GEN-DO		OCOMP		Total (76)
	1*	14	2	13	3	12	4	11	5	6	8	10	7	9	
Coordinate clause		1	2		4	3			2	3	2	3	1	1	22
Adverbial clause	1	1	2		4	2			4	3	4	2	2	1	26
Complementiser clause							1	1							2
No try, no sense or incompleteness				2	1	1	1		4	2	7	4	1	3	26

* Figures in this row indicate individual task item numbers.

Results show that only 26 (34%) instances of *failure to supply an RC* are real instances of failure in terms of inability to try or to complete a sentence, or to make the completed sentence intelligible. Most instances of this category (66%) belong to various structures of clauses as shown in the table. Such deviation from the required sentence combination indicates that when the subjects fail to produce a relative clause, they do construct intelligible sentences (50 in all) via other sentence structures, among which they are most comfortable with using coordinate and adverbial clauses (44% and 52% respectively).

In the grammaticality judgment task, five items are designed as unacceptable English sentences with each one containing an RC-related non-standard form (S3: *use of resumptive pronoun*; S4: *incorrect RC marker* – animacy; S7: *omission of obligatory RC marker*; S8: *no adjacency to head noun*; S11: *prenominal RC*). See **Appendix III** for reference to these sentences in the task and results of individual subjects' performances

sentences in this task – *The bookstore changed its name which sells science books* and *The boy stole the jewels whose father teaches law* – seem non-standard in written English even if they are communicatively acceptable in spoken English.

on this task are provided in Appendix VII. The aggregate results of the task are presented in Table 4.2.14.

Table 4.2.14 Token Counts and Percentages of Types of RC-Related Non-Standard Forms in the Grammaticality Judgment Task

Types of RC-related non-standard forms (56 subjects)	Judgment choices					
	Acceptable		Unacceptable		Not sure	
1. use of resumptive pronoun (S3)	22	39%	31	55%	3	5%
2. incorrect RC marker – animacy (S4)	4	7%	52	93%	0	
3. omission of obligatory RC marker (S7)	17	30%	38	68%	1	2%
4. no adjacency to head noun (S8)	18	32%	35	63%	3	5%
5. prenominal RC (S11)	4	7%	47	84%	5	9%

As for the five types of non-standard forms, the subjects are most sensitive to *incorrect RC marker* involving animacy and *prenominal RC* in English (a postnominal relative clause is canonical order in English) with only four subjects (7%) from each type considering them acceptable in English. This indicates that they are highly aware of the non-standard forms involving animacy and the position of a relative clause in relation to its head noun. It is interesting that most subjects, whose first languages (Mandarin, Japanese and Korean) have *prenominal RC* as a basic word order, consider it unacceptable in English with only four subjects considering the opposite and another five undecided in their judgment. The types of *use of resumptive pronoun*, *omission of obligatory RC marker* and *no adjacency to head noun* are difficult for subjects to recognise: as many as 22 (39%), 17 (30%) and 18 (32%) subjects in each task item judge them as acceptable in English. It may be tentatively said, so far as the subjects' judgment performances on the five types of non-standard forms at issue are concerned, that these three types of non-standard forms are common among L2 learners at least in this study.

In the speech and written essays of the subjects, there are 27 and 99 relative clauses respectively (see Appendix VIII). Among the 27 relative clauses in the speech, three instances (11%) involve *incorrect RC marker*, two instances (7%) involve relativisation on the topic component in a topic-comment structure, and two instances (7%) involve *omission of obligatory RC marker*. Among the 99 written relative clauses, there is only one instance (1%) for both *incorrect RC marker* and relativisation on the topic component in a topic-comment structure, and three instances (3%) of *no adjacency to the*

head noun, but there are eight instances (8%) of *omission of obligatory RC marker*. As has been shown, there are more relative clauses in the written essays than in the speech (99 versus 27), but the percentages of instances exemplifying the types of RC-related non-standard forms in the conversation data are much higher than those in written essays (26% versus 13%).

Now we consider the occurrence of the types of RC-related non-standard forms across different tasks, the overall results of which are provided in Table 4.2.15.

Table 4.2.15 Token Counts and Percentages of Types of RC-Related Non-Standard Forms across Different Tasks

Features (types of RC-related non-standard forms)	Repetition (N = 720)*		Sentence combination (N = 784)		Grammaticality judgment** (N = 280)		Conversa- tion (N = 27)	Written essays (N = 99)	Total
1. Different relativisation	114	21%	312	64%	—***	—	—	—	426
2. Failure to produce the required RC	370	69%	74	16%	—	—	—	—	446
3. Incorrect RC marker	22	4.1%	37	8%	4	7%	3	1	67
4. No adjacency to the head noun			19	3.9%	18	32%		3	40
5. Topic-comment structure	10	1.9%	7	1.4%	—	—	2	1	20
6. Resumptive words	3	0.6%	21	4.3%	22	39%			46
7. Omission of obligatory RC marker	10	1.9%	8	1.6%	17	30%	2	8	45
8. Prenominal relative clause					4	7%			4
9. Different Meaning	8	1.5%			—	—	—	—	8
10. Inaccessible head noun and others			9	1.8%	—	—			9
Total	537		489		65		7	13	1111

* Figures in parentheses indicate respectively the aggregate number of relative clauses repeated, combined, judged, and produced in naturalistic speech and writing across different tasks.

** Only the results of the five task items involving RC-related non-standard forms are presented in the table. Percentages provided in this column are on a single-item basis (see Table 4.2.14).

*** — means 'not applicable'.

Instances of *different relativisation* (features 1) and *failure to produce the required RC* (feature 2) from the repetition and sentence combination tasks constitute the overwhelming majority of all RC-related non-standard forms (79%). These two features are not relevant for the remaining tasks. Understandably, the repetition task has more instances of *failure to produce the required RC* than the sentence combination because

the former involves more demanding listening comprehension (hearing only once) and oral production while the latter involves written production with more time and chances of back-reference.

Incorrect RC marker (feature 3) is a very common type of non-standard forms which spreads across all tasks. The majority of non-standard forms of this type come from the repetition task (22 non-standard forms by 16 subjects) and the sentence combination task (37 non-standard forms by 22 subjects). There are only three non-standard forms of this type found in the conversation data, and only one in the subjects' writing. There are also four exemplifying instances of this feature in the grammaticality judgment task. Specifically, this feature involves the misuse of case or animacy for relative pronouns, or the replacement of an RC marker with a word other than a relative pronoun (see **Appendices V and VI** for illustrative examples). In terms of misused relative pronouns relating to animacy, four subjects failed to recognise such non-standard forms in the grammaticality judgment task, three subjects made five such non-standard forms in the repetition task and nine subjects made eleven such non-standard forms in the sentence combination task. Excluding the comprehension-oriented grammaticality task, the subjects' non-standard forms of this feature in various production tasks involve mostly in case (40 instances), followed by animacy (16 instances) and only a few in the use of other words in lieu of relative pronouns (7 instances). In a word, *incorrect RC marker* is a common type of non-standard forms, and when the subjects use such forms in their interlanguage, they tend to use more of those concerning case than those concerning animacy.

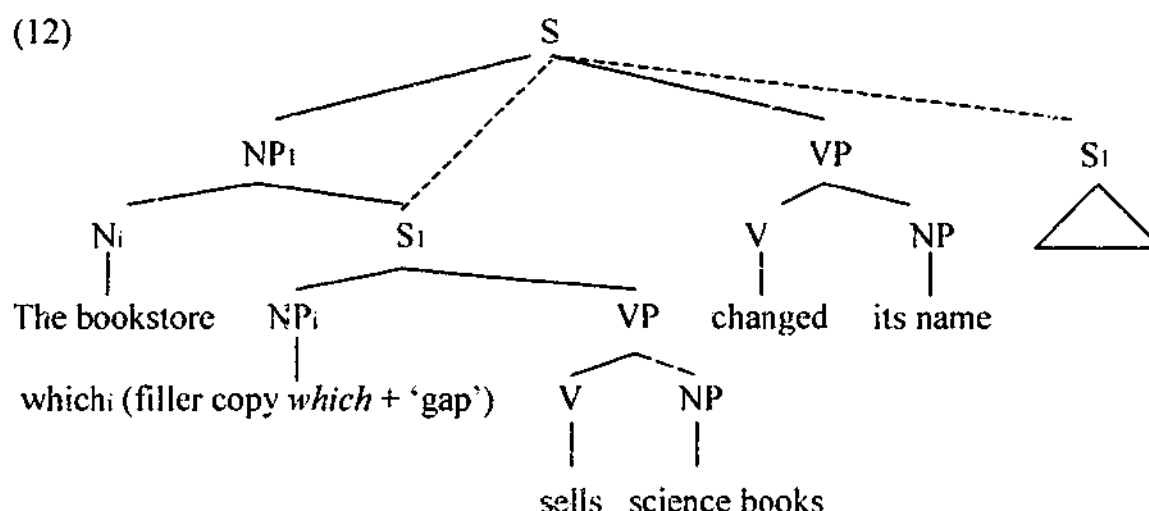
No adjacency to the head noun (feature 4) seems to be confined to the sentence combination and grammaticality judgment tasks and the subjects' written essays. Ten subjects produced instances of this feature in the sentence combination task, so did three subjects in their written essays. There were also about one third of the subjects who judged the instance of this feature (S8) as acceptable in English in the grammaticality judgment task. There were no such non-standard forms used in the repetition task and the conversation data. As has been discussed above, the tokens of this feature are acceptable

in informal spoken English, but usually not in formal written English. Even if this is the case, in both the repetition and speech of the subjects, there was no exemplifying example of this feature; whereas in the sentence combination and grammaticality judgment tasks they tended to produce and accept instances of *no adjacency to the head noun* to a considerable extent. The subjects' performances relating to this feature across different tasks with the absence of such a feature in their oral production may be due to a function of the processing load carried by the relevant structures as well as the nature of the various tasks.

As compared with the sentence combination and grammaticality judgment tasks and even with the naturalistic written essays, the repetition and conversation tasks involve more demanding constraints on language comprehension and production in terms of on-line processing. With a relative clause involving *no adjacency to the head noun*, the size of its filler-gap domain is larger than an RC which is adjacent to its head noun because the filler-gap domain of the former will include additional irrelevant nodes to process, hence more processing load. Consider the two sentences in (11).

- (11) a. *The bookstore_i s[which_i O_i sells science books]* changed its name.
 b. *The bookstore_i changed its name s[which_i O_i sells science books].*

The italicised part of the sentence in 11a and 11b indicates the size of their respective filler-gap domains. Obviously, the filler-gap domain in 11a is optimised because it is minimised in terms of the set of nodes to be processed within the domain, whereas 11b has a much larger filler-gap domain with more nodes to be processed within the domain. The filler-gap domains of 11a and 11b are shown in the tree diagram in (12).



(NP₁ = mother of filler; N_i = filler; NP_i = gap. cf. Hawkins 1999: 254)

From the tree diagram of (12), it can be seen that the filler-gap domain of 11a contains 5 nodes (N_i, NP₁, V, VP, S₁) for processing, but the filler-gap domain of 11b includes both that of 11a and an extra set of nodes (V, VP, NP, S). This extra set increases processing load within the filler-gap domain not only by adding more nodes to process, but also by delaying gap identification with an interrupting constituent structure *changed its name*. In the linear on-line parse of 11b, *the bookstore* is first interpreted as the subject of *changed* and is not interpreted as a filler until it reaches the gap site. hence involving more complex processing. In view of all this, it is, therefore, suggested that since both the repetition and conversation tasks involve time- and/or memory-constrained on-line processing, it is unlikely that the subjects will resort to a more processing loaded structure (that of *no adjacency to the head noun*) than a less loaded one (that of adjacency to the head noun) in expressing the same meaning.

Topic-comment structure (feature 5) in which the topic is relativised on, is not really a non-standard form in English even though it is a rare phenomenon in English which is predominantly a subject-prominent language (cf. Li and Thompson 1976). It is considered as a type of non-standard forms, however, in the repetition and sentence combination tasks, because such topic-comment structures are considered as deviating from the norms set in the tasks (stimulus sentences for the former and predicted RC embedding as required for the latter). No item with this feature was included in the

grammaticality judgment task, and there were only three instances of this structure found in the subjects' speech and writing. Even in the repetition and sentence combination tasks, this type of non-standard forms is only characteristic of a couple of individuals: seven subjects produced 17 instances of this structure, eleven of which were used by only two subjects.

Resumptive words (feature 6) include both nouns (3) and pronouns (43) which were redundantly used in the relative clause; the 46 resumptive words were used for all the relativised positions but SU on the Accessibility Hierarchy. This type of non-standard forms occurred in the tasks of repetition, sentence combination and grammaticality judgment, but not in the subjects' speech and written essays. Although there were only three such non-standard forms produced by three subjects in the repetition task, 13 subjects in the sentence combination task made 21 such instances, and 22 subjects treated this feature as acceptable in English in the grammaticality judgment task (with three more subjects undecided in this issue). Therefore, it is still a common type of RC-related non-standard forms among the subjects particularly in the latter two tasks.

Omission of obligatory RC marker (feature 7), like *incorrect RC marker*, is another common type of RC-related non-standard forms across all the tasks. Ten subjects produced 10 instances of this feature in the repetition task, four subjects produced 8 such instances in the sentence combination task, 17 subjects failed to recognise such a non-standard form in the grammaticality judgment task, and 9 subjects produced 10 such instances in the naturalistic speech and writing. Omitted obligatory RC markers usually include those that are obligatorily required for relativised SU and GEN positions. Results of this study show that there are no instances of this category on GEN position, and all the 18 instances in the repetition and sentence combination tasks involve relativised SU position.

Features 8, 9 and 10 are very uncommon task-specific types of non-standard forms. As for *prenominal relative clause*, there are only four subjects who failed to recognise it as a non-standard form in English in the grammaticality judgment task (with other five

subjects undecided on this issue), and no non-standard form of this type was found in any other task of the study. This suggests that the subjects in this study are well aware of the basic word order of relative clauses in English, i.e. postnominal rather than prenominal relative clauses. The feature of *different meaning* applies only to the repetition task when subjects retained the original relativised position but not the original words of the stimulus sentences in their repetition, and there were only eight instances of this feature in this task. Finally, the feature of *inaccessible head noun and others* is specific only to the sentence combination task in which six instances of relative clauses with inaccessible head nouns, two instances of double relativisation without a matrix clause, and one instance of omission of the subject in a relative clause were found.

To sum up, RC-related non-standard forms vary in types and quantities according to different types of tasks performed. The majority of non-standard forms come from *different relativisation* and *failure to produce the required RC*, which are confined to the repetition and sentence combination tasks with the former having more instances of *failure to produced the required RC* and the latter more of *different relativisation*. *Incorrect RC marker* and *omission of obligatory RC marker* are very common types of non-standard forms across different types of interlanguage data. *No adjacency to the head noun*, and *resumptive words* are also fairly common types of non-standard forms in the subjects' interlanguage. The former is not found in the more demanding on-line processing tasks such as repetition and conversation, and the latter is not found in the naturalistic data. *Topic-comment structure*, though found across different types of interlanguage data, is a feature characteristic of only a few individuals. *prenominal relative clause*, *different meaning*, and *inaccessible head noun and others* are uncommon types of non-standard forms, which are characteristic of infrequency, task specificity and/or idiosyncrasy.

4.2.5.2. RC-Related Non-Standard Forms via English Proficiency Levels

RC-related non-standard forms are analysed and discussed in relation to the factor of English proficiency levels of the subjects. Results in terms of aggregate counts are provided in Table 4.2.16.⁴⁹

Table 4.2.16 Token Counts and Percentages of Types of RC-Related Non-Standard Forms across Different Tasks by Levels

Features (types of RC-related non-standard forms)		English Proficiency Levels					
		Low (N = 20; N = 20)		Mid (N = 20; N = 20)		High (N = 20; N = 16)	
1. Different relativisation	A*	39	22%	35	20%	40	22%
	B	101	58%	105	65%	106	69%
2. Failure to produce the required RC	A	144	80%	139	78%	87	49%
	B	31	18%	32	20%	13	8.4%
3. Incorrect RC marker	A	5		5	3%	12	7%
	B	16	9.2%	14	8.6%	7	4.5%
	C	1		2		1	
4. No adjacency to the head noun	B	7	4%	1	0.6%	11	7.1%
	C	7	35%	7	35%	4	25%
5. Topic-comment structure	A			6		4	
	B					7	4.5%
6. Resumptive words	A	1		2			
	B	9	5.2%	6	3.7%	6	3.9%
	C	7	35%	8	40%	7	44%
7. Omission of obligatory RC marker	A	6		3		1	
	B	5		3			
	C	5	25%	7	35%	5	31%
8. Prenominal relative clause	C	3		1			
9. Different meaning	A	1		2		5	
10. Inaccessible head noun and other errors	B	4	2.3%	1	0.6%	4	2.6%

* A = repetition task; B = sentence combination task; and C = grammaticality judgment task.

** The first figure refers to the number of participants in the repetition task, and the second figure to the number of participants in the sentence combination and grammaticality tasks.

*** Results are presented in percentages, where necessary, as well as actual counts. In the repetition and sentence combination tasks, all percentages are calculated against the total number of non-standard forms a group produced across all features, whilst in the grammaticality judgment task, percentages are counted against the total number of performances on a single-task-item basis.

With regard to English proficiency levels of the subjects in the repetition task, the high-level group gained overall better results compared to the low- and mid-level groups. Recall (from 4.2.1.) that the high-level group scored much better in instances of correct repetition (40%) than the low- and mid-level groups (19% and 23% respectively);

⁴⁹ Due to the small number of non-standard forms found in subjects' speech and written essays (20 in all), the conversation and written essays are not listed in this table and subsequent tables in this section. However, the non-standard forms therein may be referred to, where necessary, in subsequent analysis and discussion.

consequently, this group had overall lower rates of non-standard forms. As for the two major types of non-standard forms, while the results are similar in *different relativisation* among the three groups, the high-level group, significantly, produced much fewer non-standard forms (49%) in *failure to produce the required RC* than the low- and mid-level groups (80% and 78% respectively). As regards the less common types of non-standard forms in this task, there is not much difference in *resumptive words* between the three groups, the high-level group has more instances in *incorrect RC marker* and *different meaning* and fewer instances in *omission of obligatory RC marker* than the other two groups. The occurrence of *topic-comment structure* is confined only to the mid- and high-level groups. In terms of the number of instances of correct repetition (45; 54; 95), the distribution of types of non-standard forms' (analysed above), and the total number of such forms produced (196; 192; 149) for low-, mid- and high-level groups respectively, it seems that the subjects' English proficiency levels do account for their interlanguage performances. In a word, in the repetition task, the high-level group performed significantly better than the mid-level group, which, in turn, performed better, though marginally, than the low-level group.

In the sentence combination task, on the other hand, the overall results seem to indicate that the high-level group performed no better than the other two groups either in *predicted RC embedding* or in types of non-standard forms. Recall (from 4.2.2.) that in instances of *predicted RC embedding*, the mid-level group scored marginally better (47%) than the low-level group (43%), which, in turn scored marginally better than the high-level group (40%). Consequently, the mid-level group (162 non-standard forms; 8.1 non-standard forms per person) also had slightly lower rates of non-standard forms than the low- (173; 8.7 non-standard forms per person) and the high-level group (154; 9.6 non-standard forms per person). However, an examination of the types of non-standard forms produced by the three groups shows another picture. The high-level group made more RC-related non-standard forms (in *different relativisation* and *topic-comment structure*) that are grammatical than the low- and mid-level groups: 113 (73%) versus 101 (58%) and 105 (65%), or 7.1 versus 5.1 and 5.3 non-standard forms per person respectively. In other words, with regard to non-standard forms that are ungrammatical including those in

failure to produce the required RC, the high-level group performed better than the low- and mid-level groups: 2.5 versus 3.6 and 2.8 non-standard forms per person respectively. A further examination shows that among the 26 instances of *no try*, *incomplete* and *nonsense chunks* in *failure to produce the required RC*, the high-level group made fewer such non-standard forms (4; 2.6%) than the low- and mid-level groups (11 for both groups; or 6.4% and 6.8% respectively). Specifically, the high-level group produced all the instances in *topic-comment structure*, no instances in *omission of obligatory RC marker*, and fewer instance in *incorrect RC marker* and more instances in *no adjacency to the head noun* than the other two groups. If all those grammatically constructed relative clauses are treated as correct instances via a less rigid norm than originally set for the task, the high-level group will have higher rates of correct instances of sentence combination than the low- and mid-level groups (90% versus 79% and 84% respectively). Based on the analysis of non-standard forms, we can therefore conclude that, generally, the high-level group performed better than the mid-level group, which in turn performed better than the low-level group even though the level difference in this task is not as striking as that in the repetition task.

In the grammaticality judgment task, the differences between the three groups in their judgment of RC-related non-standard forms are not as obvious as those in the above two tasks possibly due to lack of sufficient data. As is shown in table 4.2.16, the group differences occur only by one or a couple of tokens. Nevertheless, the overall results show that the high-level group made marginally fewer non-standard forms in their judgment than the low- and mid-level groups (17 versus 23, 25 non-standard forms, or 21% versus 23% and 25% respectively). Recall (from 4.2.3.) that in judging relative clauses with different relativised positions, the high-level group also performed marginally better than the low- and mid-level groups (56 versus 65 and 60 correct instances of judgment, or 70% versus 65% and 60% respectively). In sum, the high-level group performed consistently better than the other two groups in all the three tasks albeit to varying degree; however, the mid-level group performed marginally better than the low-level group in the previous two tasks but not so in this one

4.2.5.3. RC-Related Non-Standard Forms via First Languages

RC-related non-standard forms in different tasks are analysed and discussed in relation to the factor of the subjects' first languages. Results in terms of aggregate counts are provided in Table 4.2.17.

Table 4.2.17 Token Counts and Percentages of Types of RC-Related Non-Standard Forms across Different Tasks by L1s

Features (types of RC-related non-standard forms)		First Languages**							
		M A: 18 B, C: 18	J A: 10 B, C: 10	K A: 9 B, C: 7	V A: 7 B, C: 6	T A: 7 B, C: 7	In A: 3 B, C: 3	R A: 5 B, C: 4	B A: 1 B, C: 1
1. Different relativisation	A*	52 35%	16 17%	18 20%	9 14%	11 19%	5 20%	3 6.3%	
	B	99 66%	33 53%	31 60%	39 89%	44 54%	26 70%	34 62%	6
2. Failure to produce the required RC	A	86 57%	73 77%	63 69%	47 73%	42 72%	16 64%	38 79%	5
	B	21 14%	20 32%	5 10%	3 6.8%	14 17%	2 5.4%	9 16%	2
3. Incorrect RC marker	A	5 3.3%	3 3.2%	3 3.3%	1	3 5.2%	1	6 13%	
	B	6 4%	3 4.8%	8 16%		5 6.2%	6 16%	9 16%	
	C	1		1				2	
4. No-adjacency to head noun	B	9 6%	1	1	1	4 5%		1	2
	C	7	2	4	2			2	1
5. Topic-comment structure	A			6 6.5%	4.7%	1			
	B					6 7.4%			
6. Resumptive words	A	1	1	1					
	B	6 4%		4 7.8%		7 8.6%	2 5.4%	2 3.6%	
	C	8	2	1	2	3	1	4	1
7. Omission of obligatory RC marker	A	4 2.7%	2 2.1%	1	1		1	1	
	B	3 2%	5 8.1%						
	C	5	5	2		3		1	1
8. Prenominal relative clause	C	2	1	1					
9. Different meaning	A	2			3 4.7%	1	2 8%		
10. Inaccessible head noun and others	B	5 3.4%		2 3.9%		1	1		
Total	A	150	95	92	64	58	25	48	5
	B	149	62	51	44	81	37	55	10
	C	23	10	9	4	6	1	9	3

* A = repetition task; B = sentence combination task; and C = grammaticality judgment task.

** First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish, French and Italian. The figures after A, B and C refer to the number of participants in the tasks of repetition, sentence combination and grammaticality judgement respectively.

*** Results are presented in percentages, where necessary, as well as actual counts. In the repetition and sentence combination tasks, all percentages are calculated against the total number of non-standard forms a group made across all features, whilst in the grammaticality judgment task, percentages are counted against the total number of performances on a single-task-item basis.

In the repetition task, the average rate of producing non-standard forms is the lowest for the Mandarin-, Thai- and Indonesian-speaking groups (all with a rate of 8.3 non-standard forms per person), followed by the Vietnamese- (9.1), Japanese- (9.5) and Romance-speaking (9.6) groups with the Korean-speaking group having the highest (10.2).

However, regardless of those non-standard but grammatical forms (in *different relativisation*, *topic-comment structure* and *different meaning*), an examination of those non-standard and ungrammatical forms made by each group shows a different picture. It is the Mandarin-speaking group that has the lowest rate of ungrammatical non-standard forms per person (5.3), followed by the Indonesian- (6), Thai- (6.4), Vietnamese- (7), Korean- (7.6), and Japanese-speaking (7.9) groups with the Romance-speaking group having the highest rate (9). It is interesting that the difference between the Mandarin- and Romance-speaking groups in making ungrammatical non-standard forms is as huge as about four instances per person on average. A further examination shows that the Mandarin-speaking group has the highest percentage in *different relativisation* (35%) and the lowest percentage in *failure to produce the required RC* (57%) amongst all groups. The Romance-speaking group, on the other hand, has the lowest percentage in *different relativisation* (6.3%), and a higher percentage not only in *failure to produce the required RC* (79%), but in *incorrect RC marker* (13%) than any other group. A closer examination reveals that large variation within the Romance-speaking group contributes greatly to the performances of the group as a whole. Three subjects in the group whose first language is Spanish made 33 of the 38 instances in *failure to produce the required RC*, while the other two subjects, very fluent interlanguage speakers whose first languages are French and Italian respectively, used all the six incorrect RC markers of the group. The performances of the three Spanish-speaking subjects may well be due to the fact that their English proficiency levels are relatively low (low level: 2; mid level: 1), and, particularly, they are not good at listening and speaking (exemplified in the conversation and repetition tasks). For the performances of the French and Italian speakers, it is also tentatively suggested that the more proficient and confident the two subjects feel with their English speech, the more their attention is directed to meaning and fluency rather than to form and accuracy.

In the sentence combination task, it is the Japanese-speaking group that has the lowest rate of making non-standard forms (6.2 per person), followed by the Korean- (7.3), Vietnamese- (7.3), and Mandarin-speaking (8.3) groups with the Thai- (11.6), Indonesian- (12.3) and Romance-speaking (13.8) groups all exceeding ten such forms per

person. If only ungrammatical non-standard forms are taken into account, i.e. those in *different relativisation* and *topic-comment structure* are disregarded, the results are somewhat different. The Vietnamese-speaking group, surprisingly, has a much lower rate of ungrammatical non-standard forms (0.7 per person) than any other group. The Mandarin-speaking group comes out second (2.8) followed by the Japanese- (2.9), Indonesian- (3.7), and Thai-speaking (4.4) groups, still with the Romance-speaking group having the highest rate (5.3).

Recall (from 4.1.4) that we suggest for the Vietnamese-speaking group's better performance on English question formation to be a function of their high motivation, special learning conditions and great attention to form. Recall also (from (10) in 4.2.2) that their attention to form rather than to meaning is exemplified in three instances (more than any other group), in which grammatically formed relative clauses contain slight alterations to meaning. Again, this tentative account seems to be further justified in the present task in which they are not only amongst the groups that performed better in *predicted RC embedding*, but they far surpass all the other groups by producing the fewest ungrammatical non-standard forms.

As for the Romance-speaking group, all the four subjects contribute to the highest rates in both non-standard forms and ungrammatical non-standard forms. However, the very high percentage of instances of *incorrect RC marker* (16%) for this group results from only one subject (S1a, S1b), who had eight out of the nine such instances for the group in two rounds of data collection. Such idiosyncratic performances in relation to different types of non-standard forms can also be found amongst some other groups. For the Mandarin-speaking group, two subjects (M1 and M7) produced all the nine instances in *non-adjacency to head noun* and four out of the five instances in *inaccessible head noun*, and one subject (M4) had four out of the six instances in *resumptive words*. For the Thai-speaking group, one subject (T1) produced all the six instances in *topic-comment structure*, and for the Indonesian-speaking group, one subject (In2) produced all the two instances in *resumptive words* and four out of the six instances in *incorrect RC marker*. It seems a bit surprising that the topic-comment structure is employed by the Korean-,

Vietnamese- and Thai-speaking subjects, but not by the Mandarin- and Japanese-speaking subjects whose first languages do use such structure quite frequently particularly in the case of Mandarin.

In the grammaticality judgment task, it is the Indonesian-speaking group that scored the lowest in terms of the number of overall instances of making an incorrect judgment (1; 7%). The Vietnamese-speaking group is the second lowest (4; 13%) followed by the Thai- (6; 17%), Japanese- (10; 20%), Mandarin- (23; 26%), Korean-speaking (9; 26%) groups, once again, with the Romance-speaking group with the highest (9; 45%). All the four Romance-speaking subjects contribute, roughly equally, to the group's under-par performances. Specifically, only one subject was able to judge as unacceptable the instance of *omission of obligatory RC marker*; more interestingly, all the four subjects judged as acceptable the relative clause with a resumptive pronoun. It is indeed surprising that the Romance-speaking subjects, whose first languages (Spanish and Italian) are more compatible with English in relative clause formation, performed consistently in producing more instances of (ungrammatical) non-standard forms across the three tasks than all the Asian-language-speaking groups. There is no straightforward explanation for it. This is perhaps partly due to the performances of the low-level subject (S1a, S1b), who participated in both rounds of data collection and scored even below the average low-level subject on both occasions. Yet the other subjects in this group scored no better than the average scores achieved at their respective proficiency levels. All this at least suggests that the subjects' different L1 backgrounds do not seem to affect their L2 performances, for instance, as shown in their judgment of the relative clause with a resumptive pronoun.

4.2.6. Summary

In conclusion, the data concerning relative clause formation with reference to the Accessibility Hierarchy in various tasks in the study seem to overwhelmingly support the validity of the predictive power of the typological hierarchy. Not only do the subjects' performances in terms of the predicted relative clause formation support the predictions made by the Accessibility Hierarchy, but their deviations from the predicted in terms of

different relativisation also lend strongly to the hierarchy. The processing-based Accessibility Hierarchy (Hawkins 1999) is more convincingly accountable for interlanguage data in this study than not only other interpretations but also the approach of the Accessibility Hierarchy interpreted otherwise. Nevertheless, two issues concerning the predictive validity of the Accessibility Hierarchy do arise. The first pertains to the relativised GEN position in the data: the Accessibility Hierarchy alone can not adequately address the subjects' performances on this relativised position, which is discussed in Chapter 6 (see 6.2 for details). The second issue concerns the relative clauses used in the spontaneous interlanguage speech: the Absolutive Hypothesis (Fox 1987) seems to account a bit better for the naturalistic interlanguage speech than the Accessibility Hierarchy. In spite of these issues, the Accessibility Hierarchy interpreted from a processing perspective proves to be a valid tool in interpreting and predicting the interlanguage data of this study with respect to relative clauses.

As far as non-standard forms in various tasks are concerned, they demonstrate that L2 learners' performances on English relative clauses do vary according to the types of tasks, their English proficiency levels and first languages. Importantly, however, all these variations seem to little affect their performances in the way predicted by the Accessibility Hierarchy.

4.3. Results regarding Other Implicational Universals

This section provides an analysis and discussion of results in the grammaticality judgment task regarding testing in interlanguage data some typological universals, either revisited or newly motivated, in Hawkins (1999). The implicational hierarchies/hypotheses tested in the task include the Clause Embedding Hierarchy (4.3.1), the Subordinate Gap/No Gap Hierarchy (4.3.2), the Hierarchy for Complementisers (4.3.3), the Bridging Verb Hierarchy (4.3.4), the Head Noun Phrase Hierarchy (4.3.5), and the Valency Completeness Hypothesis (4.3.6), which is then followed with a summary (4.3.7). The subjects' performances on the task items in relation to these implicational universals are presented in **Appendix IX**.

4.3.1. Results regarding the Clause Embedding Hierarchy

The Clause Embedding Hierarchy (see 2.1.4 for details) is examined in the grammaticality judgment task in the form of *Hypothesis IV* (3.2.4), which is re-stated in (13).

- (13) **Hypothesis IV:** There is no relation between the filler-gap domains in different types of clauses as noted in the Clause Embedding Hierarchy (Hawkins 1999), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing a filler-gap domain with the gap in a infinitival phrase will not be greater than those of judging as acceptable the sentence containing a filler-gap domain with the gap in a finite subordinate clause, which, in turn, will not be greater than those of judging as acceptable the sentence containing a filler-gap domain with the gap in a complex NP.

This hypothesis was tested in two sets of task items. The first set includes S25, S29 and S33, which are presented in (14).

- (14) S25 The person that you tried to see is Harry.
S29 The person that you hoped that you would see is Harry.
S33 The student who you know the professor that taught is Harry.
(S33' The person that you thought that Bill said Harry wanted to see has gone overseas.)

The three task items (S25, S29 and S33) were tested in the first round of data collection. In the second round, however, S33' instead of S33 was used with S25 and S29 remaining the same. The overall results of the subjects' performances on these items in the two rounds of data collection are provided in Table 4.3.1.

Table 4.3.1 Token Counts and Percentages for
Task Items Relating to the *Clause Embedding Hierarchy* (1)

Data collection	Task item number	Judgment choices					
		Acceptable		Unacceptable		Not Sure	
Round one (23 subjects)	S25	18	78.3%	4	17.4%	1	4.3%
	S29	14	61%	7	30%	2	9%
	S33	6	26%	16	70%	1	4%
Round two (33 subjects)	S25	25	76%	7	21%	1	3%
	S29	10	30%	23	70%	0	
	S33'	13	39%	16	49%	4	12%

The overall results from the first round of data collection show that the subjects' performances on S25, S29 and S33 are in compliance with the Clause Embedding Hierarchy. Results via a Chi-square analysis show that the relation between the filler-gap domains in different types of clauses and the subjects' responses to them, is significant ($\chi^2 = 13.12$, $df = 2$, $p < .01$).⁵⁰ That is, significantly, S25 (containing an embedded subcategorised gap in an infinitival phrase) is regarded as more acceptable than S29 (containing such a gap in a finite subordinate clause), which is, again, more acceptable than S33 (containing such a gap in a complex NP environment). Since the null hypothesis (*Hypothesis IV*) is rejected, it can therefore be concluded that the first-round interlanguage data predominantly support the hierarchy.

In the second round, the original S33 was changed into S33' in an attempt to see whether the increased filler-gap domain in terms of extraction of the filler from the extended subordinate structure (S33') rather than from the complex NP structure (S33) also created the increased unacceptability in the subjects' judgment. Results from the second round show that as compared with the first-round instances of 'acceptable' judgment in percentage terms, while S25 is judged similarly, there is, surprisingly, a sharp drop for S29 (from 61% to 30%). However, the significant difference between S29 and S33 in the first round does not exist between S29 and S33' in the second round. In effect, the second-round instances of S33' being judged as acceptable are marginally more than those for S29, indicating that extracting the filler from the extended subordinate structure

⁵⁰ Due to the small number of 'not sure' responses, 'not sure' and 'unacceptable' responses are collapsed into one level of measurement, as opposed to the other level of measurement – 'acceptable', in running Chi-square, hence the degree of freedom of 2 instead of 4 [the formula of calculation: $(2 - 1) \times (3 - 1)$].

(S33') as compared with extracting a filler from a closer subordinate structure (S29) does not create extra processing difficulty in the subjects' judgment. Notwithstanding this, when the extracted filler comes from a more complex structural environment – the NP complex structure as in S33, the processing difficulty does increase, hence the fewer instances of regarding such a structure as acceptable.

Due to the relatively fewer exemplifying instances when the data are analysed separately in two rounds, it is not considered as adequately appropriate to examine the data further in terms of the subjects' English proficiency levels and first languages. Now we turn to the examination of the second set of task items S22 and S28, which are presented in (15).

- (15) S22 What did you wonder how they would bake?
 S28 What did you wonder how to bake?

The overall results (see **Appendix IX**) of the subjects' performances on the two items are presented in Table 4.3.2.

Table 4.3.2 **Token Counts and Percentages for**
Task Items Relating to the *Clause Embedding Hierarchy* (II)

Task item number (56 subjects)	Judgment choices					
	Acceptable		Unacceptable		Not Sure	
S22	24	43%	24	43%	8	14%
S28	26	46%	23	41%	7	13%

It is obvious from the overall results in the table that the subjects performed similarly on S22 and S28 across the three judgment choices, hence there is no need for a Chi-square test and no rejection of the null hypothesis (*Hypothesis IV*) in this set of data. Although the Clause Embedding Hierarchy is not supported, it is not disconfirmed either in the subjects' performances. The indiscriminate responses of the subjects to the two stimulus sentences may indicate that such rarely occurring structures in English are too complex (containing two gaps in an infinitival phrase and a finite subordinate clause in S28 and S22 respectively) for them to comprehend fully. Some students reflected after the task that they had seen interrogative sentences beginning with two *wh*-words such as *when*

and where, how and why, but hardly met with one *wh*-word at the beginning and another in the middle of a sentence. As has been illustrated in section 2.1.4, filler-gap domains are complex structures to process, and filler-gap domains containing another gap, which border on grammaticality, are even more complex for processing. In fact, more than half of the subjects (30) treat the two task items exactly the same way, namely, both as 'acceptable', both as 'unacceptable', or both as 'not sure'. We now examine whether factors of the subjects' English proficiency levels and first languages may affect the overall results. Results are provided in Table 4.3.3.

Table 4.3.3 Token Counts and Percentages for Task Items
Relating to the *Clause Embedding Hierarchy* (II) by Levels and L1s

Task item number	English Proficiency Levels			First Languages***							
	Low (20)**	Mid (20)	High (16)	M (18)	J (10)	K (7)	V (6)	T (7)	In (3)	R (4)	B (1)
S22 1*	10 50%	9 45%	5 31%	7 39%	4 40%	3 43%	4 67%	4 57%	1 33%	1 25%	1
2	9	8	7	9	5	3	1	2	2	1	1
3	1	3	4	2	1	1	1	1	1	2	
S28 1	9 45%	11 55%	6 38%	10 56%	5 50%	4 57%	4 67%	2 29%	1 33%	3	1
2	9	5	9	7	5	1	2	2	2	1	
3	2	4	1	1		2		3		1	

* Coded judgment choices: 1 = acceptable; 2 = unacceptable; and 3 = not sure.

** Figures within parentheses indicate the number of subjects in each group.

*** First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish and Italian

Results from the table show that disregarding those one-token differences between S22 and S28 in terms of instances of judging them as acceptable, the performances of the mid-level group and Mandarin-speaking group tend to confirm the hierarchy while the Thai-speaking group tends to disconfirm it. The differences, however, are not significant, and, overwhelmingly, the performances of all groups are consistent with the overall results.

To sum up, the Clause Embedding Hierarchy is supported by the performances of the subjects on the first set of task items pertaining to this hierarchy in the grammaticality judgment task, but not by those on the second set which involve extremely complex structures to process.

4.3.2. Results regarding the Subordinate Gap/No Gap Hierarchy

The Subordinate Gap/No Gap Hierarchy (see 2.1.4 for details) is examined in the grammaticality judgment task in the form of *Hypothesis V* (3.2.4), which is re-stated in (16).

- (16) **Hypothesis V:** There is no relation between the filler-gap domains containing an extra gap and no extra gap in a subordinate clause as noted in the Subordinate Gap/No Gap Hierarchy (Hawkins 1999), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing a filler-gap domain in which a matrix filler can be matched with a gap in a subordinate clause of complexity *n* containing another gap, will be greater than those of judging as acceptable the sentence containing a filler-gap domain in which the matrix filler can be matched with a gap in subordinate clause of complexity *n* containing no other gap.

This hypothesis was tested in task items S24 and S31, which are presented in (17).

- (17) S24 Who do you know the professor that taught?
S31 What do you regret the fact that he stole?

The overall results (see Appendix IX) of the subjects' performances on these two items are presented in Table 4.3.4.

Table 4.3.4 Token Counts and Percentages for
Task Items Relating to the *Subordinate Gap/No Gap Hierarchy*

Task item number (56 subjects)	Judgment choices					
	Acceptable		Unacceptable		Not Sure	
S24	11	20%	42	75%	3	5%
S31	26	46.4%	26	46.4%	4	7%

Recall (from 2.1.4.) that although both sentences are ungrammatical, S31 (containing one gap within its filler-gap domain), is still regarded as more grammatical than S24 (containing two gaps within its filler-gap domain), and may "even border on grammaticality" (Hawkins, 1999: 270). This was actually borne out in the subjects'

judgment of the two sentences: S31 was indeed regarded as more acceptable (46.4%) than S24 (20%). Using a Chi-square test, the relation between the types of filler-gap constructions and the subjects' overall performance is statistically significant ($\chi^2 = 7.91$, $df = 1$, $p < .01$).⁵¹ Consequently, the null hypothesis (*Hypothesis V*) is rejected, and the Subordinate Gap/No Gap Hierarchy is upheld by the interlanguage data. Again, we now examine whether the overall results are subject to the impact of factors of the subjects' English proficiency levels and first languages. Results of the examination are presented in Table 4.3.5.

Table 4.3.5 Token Counts and Percentages for Task Items
Relating to the *Subordinate Gap/No Gap Hierarchy* by Levels and L1s

Task item number	English Proficiency Levels			First Languages***							
	Low (20)**	Mid (20)	High (16)	M (18)	J (10)	K (7)	V (6)	T (7)	In (3)	R (4)	B (1)
S24 1*	6 30%	3 15%	2 13%	4 22%	5 50%			1 14%		1 25%	
2	13	15	14	13	5	6	6	5	3	3	1
3	1	2		1		1		1			
S31 1	9 45%	9 45%	8 50%	10 56%	5 50%	2 29%	5 83%	2 29%	1 33%	1 25%	
2	10	9	7	7	4	5	1	3	2	3	1
3	1	2	1	1	1			2			

* Coded judgment choices: 1 = acceptable; 2 = unacceptable; and 3 = not sure.

** Figures within parentheses indicate the number of subjects in each group.

*** First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish and Italian

Results from the table show that apart from the Japanese- and Romance-speaking groups that treated S24 and S31 indiscriminately, all the other groups regarded S31 as more acceptable than S24. All the data are, nonetheless, in conformity with the Subordinate Gap/No Gap Hierarchy. In short, the Subordinate Gap/No Gap Hierarchy proves to be thoroughly supported by the interlanguage data under examination.

4.3.3. Results regarding the Hierarchy for Complementisers

The Hierarchy for Complementisers (see 2.1.4 for details) is examined in the grammaticality judgment task in the form of *Hypothesis VI* (3.2.4), which is re-stated in (18).

⁵¹ Since it was a two-way test with $df = 1$, Yates' Correction Factor was used (Hatch and Lazaraton: 404-5).

- (18) **Hypothesis VI:** There is no relation between the filler-gap domains containing different types of complementisers as noted in the Hierarchy for Complementisers (Kluender 1992), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing *that* as acceptable will not be greater than those of judging as acceptable the sentence containing *if*, which, in turn, will not be greater than those of judging as acceptable the sentence containing *whether*.

Task items S10, S13, and S16 were used to test the hypothesis with item S20 used as a comparative sentence containing a less complex structure than the other three. The four task items are presented in (19).

- (19) S10 What did John doubt whether she would win?
 S13 What did John doubt if she would win?
 S16 What did John doubt that she would win?
 S20 What did John doubt the truth of?

The overall results (see **Appendix IX**) of the subjects' performances on these items are presented in Table 4.3.6.

Table 4.3.6 Token Counts and Percentages for
 Task Items Relating to the *Hierarchy for Complementisers*

Task item number (56 subjects)	Judgment choices					
	Acceptable		Unacceptable		Not Sure	
S10 (<i>whether</i>)	11	20%	32	58%	12	22%
S13 (<i>if</i>)	24	43.6%	24	43.6%	7	12.7%
S16 (<i>that</i>)	26	46%	23	41%	7	13%
S20	31	55.4%	13	23.2%	12	21.4%

Results of a Chi-square analysis indicate that the null hypothesis (*Hypothesis VI*) is rejected ($\chi^2 = 10.02$, $df = 2$, $p < .01$), and that the subjects' overall performances on S10, S13 and S16 significantly support the Hierarchy for Complementisers. That is, S16 (containing *that*, the least semantically loaded complementiser to be processed) is regarded as more acceptable (albeit slightly) than S13 (containing *if*, a more semantically loaded complementiser), which is, in turn, more acceptable than S10 (containing *whether*, the most semantically loaded complementiser). The processing account is further justified

in the subjects' even better performances on S20, which has a much less complex filler-gap domain in which the gap is subcategorised by the preposition *of* in a simple (interrogative) clause rather than subcategorised by a verb within a *complementiser*-clause in the above three examples. We now turn to the examination of the results on the basis of proficiency levels and first languages of the subjects to see whether these they may affect their overall performances. Results are provided in Table 4.3.7.

Table 4.3.7 Token Counts and Percentages for Task Items
Relating to the *Hierarchy for Complementisers* by Levels and L1s

Task item number	English Proficiency Levels			First Languages***							
	Low (20)**	Mid (20)	High (16)	M (18)	J (10)	K (7)	V (6)	T (7)	In (3)	R (4)	B (1)
S10 1*	7 35%	4 20%		2 11%	4 40%	1 14%	1 17%	1 17%	1 33%	1 25%	
2	11	11	10	12	6	4	2	4	2	1	1
3	2	4	6	4		2	3	1		2	
S13 1	10 50%	8 40%	6 40%	7 39%	5 50%	2 29%	2 33%	3 50%	3 100%	2 50%	
2	9	9	6	8	4	5	3	2		1	1
3	1	3	3	3	1		1	1		1	
S16 1	7 35%	11 55%	8 50%	9 50%	4 40%	2 29%	4 67%	4 57%	2 67%	1 25%	
2	11	7	5	7	5	4	2	1	1	2	1
3	2	2	3	2	1	1		2		1	
S20 1	11 55%	11 55%	9 56%	8 44%	8 80%	3 41%	5 83%	2 29%	3 100%	2 50%	
2	5	3	5	7	2	1	1	2			
3	4	6	2	3		3		3		2	1

* Coded judgment choices: 1 = acceptable; 2 = unacceptable; and 3 = not sure.

** Figures within parentheses indicate the number of subjects in each group.

*** First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish and Italian.

With regard to the subjects' English proficiency levels, results show that the mid- and high-level groups' performances are in compliance with the Hierarchy for Complementisers, whilst the low-level group's performances are not. For the low-level group, S13 containing *if* was judged to be more acceptable (50%) than S10 containing *whether* and S16 containing *that*, both of which were judged to be equally acceptable (35%). This may tentatively indicate that the low-level subjects may have difficulty understanding the semantic implications of the three complementisers, and may also have used *if* more often in their interlanguage, be it a *conditional* or *complementiser*, hence the result. Nonetheless, all the three groups treat S20, with a less complex filler-gap domain as more acceptable than (or at least as equally acceptable as) the three sentences with a complementiser-clause.

Amongst L1 groups, the performances of the Mandarin-, Korean-, Vietnamese- and Thai-speaking groups comply with the Hierarchy for Complementisers, whereas the Japanese-, Indonesian- and Romance-speaking groups do not, all judging S13 (with *if*) as more acceptable than S16 (with *that*). However, in each case the difference between S13 and S16 is just one token, which, as indicated in our previous analysis, is not considered as a valid piece of evidence for either proving or disproving a pattern of occurrence. As for the acceptable status of S20, the Japanese-, Vietnamese- and Indonesian-speaking groups all have a much higher level of acceptance over that of the other three items, while the Thai-speaking group has a relatively lower level of acceptance of S20 over that of S13 and S16.

In closing, the Hierarchy for Complementisers is generally supported by the subjects' performances in the grammaticality judgment task with the exception of the low-level group, the issue of which we will return to in the next section of this chapter.

4.3.4. Results regarding the Bridging Verb Hierarchy

The Bridging Verb Hierarchy (see 2.1.4 for details) is examined in the grammaticality judgment task in the form of *Hypothesis VII* (3.2.4), which is re-stated in (20).

- (20) **Hypothesis VII:** There is no relation between the filler-gap domains containing bridging verbs of varying semantic specificity as noted in the Bridging Verb Hierarchy (Hawkins 1999), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing a bridging verb (complex) V of semantic specificity within a filler-gap domain, will be not be greater than those of judging as acceptable the sentence containing a bridging verb (verb complex) V' with more semantic specificity than V.

This hypothesis was tested in task items S18, S21 and S26, which are presented in (21).

- (21) S18 How angry did Mary say that John was?
 S21 How angry did Mary say softly that John was?
 S26 How angry did Mary whisper that John was?

The overall results (see Appendix IX) of the subjects' performances on these items are presented in Table 4.3.8.

Table 4.3.8 Token Counts and Percentages for Task Items Relating to the *Bridging Verb Hierarchy*

Task item number (56 subjects)	Judgment choices					
	Acceptable		Unacceptable		Not Sure	
S18 (<i>say</i>)	24	43%	31	55%	1	2%
S21 (<i>say softly</i>)	14	25%	29	52%	13	23%
S26 (<i>whisper</i>)	13	23%	28	50%	15	27%

Results of a Chi-square test show that the null hypothesis (*Hypothesis VII*) is rejected ($\chi^2 = 5.35$, $df = 1$, $p < .025$), and that the subjects' performances on S18, S21 and S26 comply with the Bridging Verb Hierarchy.⁵² That is, S18 (containing *say* – the least semantically loaded bridging verb in the hierarchy) was regarded as significantly more acceptable than S21 and S26 (containing *say softly* and *whisper* respectively, both of which are more semantically loaded than *say* and were regarded as more or less equally acceptable). Again, we examine whether the subjects' English proficiency levels and first languages may affect the overall results. Results are provided in Table 4.3.9.

Table 4.3.9 Token Counts and Percentages for Task Items Relating to the *Bridging Verb Hierarchy* by Levels and L1s

Task item number	English Proficiency Levels			First Languages***							
	Low (20)**	Mid (20)	High (16)	M (18)	J (10)	K (7)	V (6)	T (7)	In (3)	R (4)	B (1)
S18 1*	6 30%	10 50%	8 50%	11 61%	3 30%	2 29%	3 50%	3 43%		2 50%	
2	14	10	7	7	7	5	3	3	3	2	1
3			1					1			
S21 1	4 20%	5 25%	5 31%	6 33%	2 20%	1 14%	4 67%	1 14%			
2	12	10	7	8	8	5	2	3	1	1	1
3	4	5	4	4		1		3	2	3	
S26 1	3 15%	7 35%	3 19%	4 22%	1 10%	3 43%	4 67%	1 14%			
2	13	8	7	9	8	3		3	2	2	1
3	4	5	6	5	1	1	2	3	1	2	

* Coded judgment choices: 1 = acceptable; 2 = unacceptable; and 3 = not sure.

** Figures within parentheses indicate the number of subjects in each group.

⁵² Because S21 and S26 are treated as semantic equivalents, they were collapsed into one level in running the Chi-square test.

*** First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish and Italian.

In view of the English proficiency levels of the subjects, the performances of all the three groups are in general conformity with the Bridging Verb Hierarchy. S18 was judged to be more acceptable than S21 and S26. All the L1 groups except the Korea- and Vietnamese-speaking groups also performed in the way predicted by the hierarchy. The seemingly disconfirming performances of the Korean- and Vietnamese-speaking groups, however, can be disregarded on an account of a one-token-only difference between S18 and S21/S26 in each case. It can be concluded, therefore, that the Bridging Verb Hierarchy is overwhelmingly supported by the performances of the subjects in the grammaticality judgment task.

4.3.5. Results regarding the Head Noun Phrase Hierarchy

The Head Noun Phrase Hierarchy (see 2.1.4 for details) is examined in the grammaticality judgment task in the form of *Hypothesis VIII* (3.2.4), which is re-stated in (22).

- (22) **Hypothesis VIII:** There is no relation between the filler-gap domains containing NPs with head nouns of varying semantic specificity as noted in the Head Noun Phrase Hierarchy (Hawkins 1999), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the sentence containing an NP with a head noun (phrase) N of semantic specificity within a filler-gap domain, will not be greater than those of judging as acceptable the sentence containing an NP with a head noun (phrase) N' of more semantic specificity than N.

This hypothesis was tested in task items S14, S17 and S23, which are presented in (23).

- (23) S14 Who did you see a picture of?
S17 Who did you see the picture of?
S23 Who did you see John's picture of?

The overall results (see Appendix IX) of the subjects' performances on these items are presented in Table 4.3.10.

Table 4.3.10 Token Counts and Percentages for
Task Items Relating to the *Head Noun Phrase Hierarchy*

Task item number (56 subjects)	Judgment choices					
	Acceptable		Unacceptable		Not Sure	
S14 (<i>a</i>)	18	32%	28	50%	10	18%
S17 (<i>the</i>)	30	53.6%	19	33.9%	7	12.5%
S23 (<i>John's</i>)	22	39%	28	50%	6	11%

Results from the table show that the subjects' overall performances do not conform to the Head Noun Phrase Hierarchy (Hawkins 1999). According to the hierarchy, an NP with an indefinite head (S14) is less semantically specific (readily permitting a gap in an *of*-complement) than one with a definite head (S17; making the filler-gap structure less acceptable), which is, in turn, less so than one with a possessive modifier (S23; the least acceptable). Grammaticality declines as the amount of semantic processing increases from S14 to S17, then to S23. It is surprising, however, that S14, the most grammatical of the three predicted by the hierarchy, was judged to be the least acceptable despite the fact that S17, in accordance with the hierarchy, was judged to be more acceptable than S23. Without running Chi-square, the null hypothesis (*Hypothesis VIII*) is thus taken not to be rejected. This disconfirming evidence against the hierarchy, though not statistically significant ($\chi^2 = 5.49$, $df = 2$, $p < .10$), may suggest that the processing account in terms of filler-gap domain alone can not address adequately the L2 learners' interlanguage behaviour under examination. Now we examine further whether the subjects' English proficiency levels and first languages may contribute to their overall performances. Results are provided in Table 4.3.11.

Table 4.3.11 Token Counts and Percentages for Task Items
Relating to the *Head Noun Phrase Hierarchy* by Levels and L1s

Task item number	English Proficiency Levels			First Languages***							
	Low (20)**	Mid (20)	High (16)	M (18)	J (10)	K (7)	V (6)	T (7)	In (3)	R (4)	B (1)
S14 1*	7 35%	6 30%	5 31%	4 22%	4 40%	4 57%	2 33%	3 43%	1 33%		
2	9	11	8	11	5	2	3	1	2	3	1
3	4	3	3	3	1	1	1	3		1	
S17 1	11 55%	14 70%	5 31%	7 39%	8 80%	4 57%	3 50%	5 71%	2 67%	1 25%	
2	7	5	7	8	2	3	1	1	1	2	1
3	2	1	4	3			2	1		1	
S23 1	10 50%	8 40%	4 25%	7 39%	5 50%	4 57%	1 17%	2 29%	1 33%	1 25%	1
2	9	11	8	9	4	3	5	2	2	3	
3	1	1	4	2	1			3			

* Coded judgment choices: 1 = acceptable; 2 = unacceptable; and 3 = not sure.

** Figures within parentheses indicate the number of subjects in each group.

*** First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish and Italian.

From the table, it is clear that apart from the performances of both the high-level group and the Korean-speaking group who judged the three sentences roughly equally in terms of acceptability, the performances of all the other groups are consistent with the overall results analysed above. Although it seems apparent that the subjects' English-learning background may have an impact on their performances,⁵³ it is more likely that the semantic complexity of the English articles contributes greatly to the counterexample, which is discussed in chapter six (6.2). In a word, whatever may be the reason, the Head Noun Phrase Hierarchy is not supported by the interlanguage data under examination.

4.3.6. Results regarding the Valency Completeness Hypothesis

The Valency Completeness Hypothesis (see 2.1.4 for details) is examined in the grammaticality judgment task in the form of *Hypothesis IX* (3.2.4), which is re-stated in (24).

- (24) **Hypothesis IX:** There is no relation between filler-gap domains that are valency complete or valency incomplete as noted in the Valency Completeness Hypothesis (Hawkins 1999), and the ESL learners' overall performance in the grammaticality judgment task. Or significantly, the subjects' instances of judging as acceptable the

⁵³ In many Asian countries, grammar/structure-oriented English instruction is still the norm; with such an approach, there tends to be an emphasis on an NP with a definite head before an *of*-phrase. The point is made on the basis of my personal experience of learning and teaching English as a second language.

sentence containing a filler-gap domain that includes the subcategorisors for all phrases within the domain that contains the gap, will not be greater than those of judging as acceptable the sentence containing a filler-gap domain that does not include all the subcategorisors for all phrases within the domain that contains the gap.

This hypothesis was tested in three sets of stimulus sentences. The first set includes S27 and S34, the second S30 and S35, and the third S32 and S36, all of which are presented in (25).

- (25) a. S27 Who did it surprise Sue that Mary disliked?
 S34 Who did that Mary disliked surprise Sue?
- b. S30 What did the title of amuse John?
 S35 What did John read the title of?
- c. S32 What did to read fascinate Sue?
 S36 What did it fascinate Sue to read?

The overall results (see **Appendix IX**) of the subjects' performances on these items are presented in Table 4.3.12.

Table 4.3.12 **Token Counts and Percentages for**
 Task Items Relating to the *Valency Completeness Hypothesis*

Task item number (56 subjects)		Judgment choices					
		Acceptable		Unacceptable		Not Sure	
a.	S27	9	16%	35	63%	12	21%
	S34	14	25%	33	59%	9	16%
b.	S30	15	27%	37	66%	4	7%
	S35	35	62.5%	15	26.8%	6	10.7%
c.	S32	7	12.5%	48	85.7%	1	1.8%
	S36	33	59%	20	36%	3	5%

Hawkins' processing-motivated hypothesis – the Valency Completeness Hypothesis states that “The human processor prefers FGDs (filler-gap domains) to include the subcategorizers for all phrases within the domain that contain the gap” (Hawkins 1999: 278). Results show that the subjects performed on the last two sets (25b and 25c) in the way predicted by the Valency Completeness Hypothesis, but that their performance on

the first set (25a) runs counter to the prediction of the hypothesis. As a result, the null hypothesis (*Hypothesis IX*) can not be rejected in the first set of data, but is, in fact rejected in the other two sets (25b: $\chi^2 = 13.04$, $df = 1$, $p < .001$; 25c: $\chi^2 = 13.04$, $df = 1$, $p < .001$). The Valency Completeness Hypothesis is thus supported in the last two sets of interlanguage data, but not in the first.

In the second set (25b), a subject-object asymmetry, in which the filler extracted from the object (S35) is grammatical and that extracted from the subject (S30) is not, can be accounted for by this processing motivation. In the filler-gap domain of both S30 and S35, the filler *what* proceeds to the subcategoriser of the gap *of*. But S35 has a filler-gap domain within which the subcategorisers *read* and *of* for all the phrases containing the gap are structurally connected and will have been processed prior to gap identification, hence processing is facilitated in this case. Whereas in S30, the filler-gap domain does not include *amuse*, and thus the filler-gap domain – *What did the title of* – is valency incomplete, which results in processing hindrance. Results show that the instances of the subjects' judging S35 as acceptable, are significantly more than those for S30 (35 versus 15; or 62.5% versus 27%), which lends support to the hypothesis of Valency Completeness of Hawkins (1999).

The third set (25c), in which S36 is grammatical and S32 is not, can be similarly explained. The path from the filler to the subcategoriser of the gap proceeds from *what* to *read* in both sentences. But the verb that subcategorises for the containing infinitival complement *to read*, namely, *fascinate*, stands outside the filler-gap domain in S32. On the other hand, in the extraposition structure of S36, *fascinate* stands within the filler-gap domain, hence the valency is complete, and the complete valency of all the subcategorisers within the filler-gap domain will facilitate rather than hinder gap identification. The performances of the subjects on these two items appear to uphold this processing hypothesis: instances of their judging S36 as acceptable are significantly more than those for S32 (33 versus 7; or 59% versus 12.5%).

In the first set (25a), again, the path from the filler to the subcategoriser of the gap proceeds from *who* to *disliked* in both sentences. Yet the verb - *surprise* - which subcategorises for the sentential complement *that Mary disliked*, stands within the filler-gap domain in the extraposition structure S27, but outside the domain in S34. The subjects' performances on these two items, however, do not accord with the processing hypothesis of Valency Completeness. That is, the Hypothesis predicts that S27 would be more acceptable than S34, however, the reverse is the case (S27: 9 or 16%; S34: 14 or 25%). Although the reverse case is not statistically significant ($\chi^2 = 0.88$, $df = 1$, $p > .10$), it may indicate that sentential subjects involve a high degree of processing difficulty for L2 learners, and that this processing difficulty may contribute to their inability to discriminate between the two sentences containing sentential subjects. In fact, thirty-one (55%) out of the 56 subjects judged the two items either as both being acceptable (9, 16%) or as both being unacceptable (22, 39%). Regardless of Valency Completeness within the filler-gap domain, the size of the filler-gap domain in S27 is much larger than that in S34. The processing difficulty for sentential subjects in combination with the larger filler-gap domain in S27 may contribute to its relatively lower degree of acceptability than that for S34. A further comparison between S27 and S36 may help shed some light on this case. Results of comparison are provided in Table 4.3.13.

Table 4.3.13 **Token Counts and Percentages for Task Items**
Relating to the *Valency Completeness Hypothesis* (for Comparison)

Task item number (56 subjects)	Judgment choices					
	Acceptable		Unacceptable		Not Sure	
S27	9	16%	35	63%	12	21%
S36	33	59%	20	36%	3	5%

S27 and S36 are both valency complete, but the former contains a sentential complement and the latter an infinitival complement. The percentage of predicted response for S36 (59%) is much higher than that for S27 (16%) and a Chi-square test shows that the relation between the valency complete filler-gap domains across different clause types and the overall performance of the subjects is highly statistically significant ($\chi^2 = 20.15$, $df = 1$, $p < .001$). As seen above, the subjects can not discriminate between S27 and S34 in terms of valency completeness, both of which contain sentential complements.

However, when the sentential complements are converted into infinitival complements as in S32 and S36, valency completeness does come into play in the subjects' judgment of the two sentences (see Table 4.3.12). It seems that the Clause Embedding Hierarchy, in which infinitival phrases are more hospitable to gaps than finite subordinate clauses, can best account for the difference between S27 and S36 (see Table 4.3.13). Now we examine whether the subjects' English proficiency levels and first languages may have an effect on their performances on the items related to the Valency Completeness. Results are provided in Table 4.3.14.

Table 4.3.14 Token Counts and Percentages for Task Items
Relating to the *Valency Completeness Hypothesis* by Levels and L1s

Task item number	English Proficiency Levels			First Languages***							
	Low (20)**	Mid (20)	High (16)	M (18)	J (10)	K (7)	V (6)	T (7)	In (3)	R (4)	B (1)
a. S27 1*	2 10%	6 30%	1 6%	4 22%	1 10%	3 43%		1 14%			
2	16	10	9	10	8	4	5	3	2	2	1
3	2	4	6	4	1		1	3	1	2	
S34 1	5 25%	6 30%	3 19%	5 28%	1 10%	3 43%	2 33%	2 29%		1 25%	
2	14	12	7	8	9	3	4	3	3	3	
3	1	2	6	5		1		2			1
b. S30 1	4 20%	6 30%	5 31%	4 22%	1 10%	3 43%	1 17%	4 57%	1 33%		1
2	16	11	10	13	9	2	5	2	2	4	
3		3	1	1		2		1			
S35 1	15 75%	10 50%	10 63%	14 78%	7 70%	4 57%	6 100%	2 29%		1 25%	1
2	3	8	4	4	1	2		2	3	3	
3	2	2	2		2	1		3			
c. S32 1	2 10%	4 20%	1 6%	2 11%	1 10%			2 29%	1 33%	1 25%	
2	18	16	14	15	9	7	6	5	2	3	1
3			1	1							
S36 1	8 40%	15 75%	10 63%	12 67%	6 60%	4 57%	2 33%	4 57%	1 33%	3 75%	1
2	10	4	6	6	3	2	4	3	1	1	
3	2	1			1	1			1		

* Coded judgment choices: 1 = acceptable; 2 = unacceptable; and 3 = not sure.

** Figures within parentheses indicate the number of subjects in each group.

*** First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish and Italian

Results from the table show that differences can be found in some cases. In the first set (a: S27 and S34), the mid-level group Mandarin- and Korean-speaking groups have higher percentages for instances of judging sentential subjects as acceptable, while the high-level group and the Japanese-, Indonesian- and Romance-speaking groups tend to have low percentages. In the second set (b: S30 and S35) regarding subject-object asymmetry, the Thai-speaking group has more instances judging S30 (with incomplete valency within the filler-gap domain) as acceptable than those for S35 (with complete valency), which is against the Valency Completeness Hypothesis. Nevertheless, results in

terms of the subjects' English proficiency levels and their first languages are consistent with the overall results analysed above.

4.3.7. Summary

In this section (4.3.), six of the implicational hierarchies/hypotheses revisited or motivated by Hawkins (1999) are tested against interlanguage data in the grammaticality judgment task in the study. The overall results tend to lend support to these universals as valid predictors for interlanguage performances with one significant exception, specifically. The Subordinate Gap/No Gap Hierarchy (4.3.2), the Hierarchy for Complementisers (4.3.3), and the Bridging Verb Hierarchy (4.3.4) are fully supported by the subjects' performances, the Clause Embedding Hierarchy (4.3.1) and the Valency Completeness Hypothesis (4.3.6) are mostly supported, but the Head Noun Phrase Hierarchy (4.3.5) is not supported.

In the case of the Clause Embedding Hierarchy, it is supported by two sets of comparison – S25, S29 & S33 (Table 4.3.1) and S27 & S36 (Table 4.3.13), but another set of comparison – S22 & S28 (Table 4.3.2) seems indiscriminate. Similarly for the Valency Completeness Hypothesis, which is supported by two sets of comparison – S30 & S35, and S32 & S36 (Table 4.3.12), another set of comparison – S27 & S34 is also indiscriminate. However, in both cases of indiscriminate performances of the subjects, the sentences involved are far more complex in structure than those in the other sets of comparison. That is, for the Clause Embedding Hierarchy, the indiscriminate set involves structures of *wh*-islands which contain two gaps in the complementiser phrases (both finite and nonfinite), while all the other sentences involve no such structure but structure that contains one gap in infinitival phrase, finite clause and complex NP respectively. For the Valency Completeness Hypothesis, the indiscriminate set involves filler-gap domains in sentential complements as subjects in both sentences, whereas all the other sentences pertain to less complex structures such as filler-gap domains in prepositional complements and infinitival complements respectively. This suggests that these universals hold in the interlanguage of L2 learners to the extent to which the complexity of the English structures they encounter is within their scope of comprehension. In other

words, when the English structures such as *wh*-islands and sentential subjects are too complex for L2 learners to comprehend, it becomes unlikely for them to recognise the minute differences that reside in the structures.

In the case of the exception to the Head Noun Phrase Hierarchy in the subjects' performances, we can only say at this phase that this universal does not hold for the interlanguage data under examination. The reason for the exception is similarly suggested as being a function of the complexity of the use of English articles and the level of competence of the L2 learners in this study particularly in their acquisition of the English articles. The English articles are very complex in usage, especially for L2 learners whose first languages (such as Chinese) do not have such a category, and the subjects may have not yet reached the stage in which they can comfortably use definite and indefinite articles in different types of contexts. As a result, they could not decipher the implications carried by the use of the definite and indefinite articles within a complex filler-gap domain (as provided in the stimulus sentences used for examining the Head Noun Phrase Hierarchy).

The subjects' English proficiency levels and first languages usually do not affect the overall results. Although there are a couple of disconfirming cases in these respects such as the low-level group's performances pertaining to the Hierarchy for Complementisers and the Thai-speaking group's judgment on sentences with subject-object asymmetry in relation to the Valency Completeness Hypothesis, all groups overwhelmingly conform to the universals under testing.

In sum, the implicational hierarchies and hypotheses of Hawkins (1999) tested in this study are predominantly supported by the interlanguage data in this study. This not only confirms the validity of predictive power of these universals for second language acquisition but suggests that the universal constraints from the processing perspective play a substantial role in second language acquisition as well as in first language acquisition.

4.4. Results regarding Interlanguage features in Naturalistic Data

In this section, interlanguage features of the naturalistic data are examined and analysed with reference to the results in previous three sections. The examination and analysis are presented first with respect to the overall results gained from both the speech and written essays of the subjects (4.4.1). The results are also examined in terms of different English proficiency level groups (4.4.2) and in terms of different L1 groups (4.4.3). Section 4.4.4 provides some summarising remarks.

4.4.1. Overall Results from the Conversation and Written Essays

Interlanguage features relating to some of the analytically and deductively analysed results were further examined, in the speech and written essays of the subjects. The features examined here include filler-gap constructions, passive voice, coordinate and subordinate structures, topic-comment structure, the non-standard use of subject-verb agreement and noun number, and omission of required indefinite articles. The performances of all the individual subjects in the conversation task and the written essays in terms of token counts and occurrence rates for the above grammatical features are presented in **Appendices X and XI** respectively, the overall results of which are provided in Table 4.4.1.

Table 4.4.1 **Counts and Rates for the Use of Different Grammatical Features in Naturalistic Speech (S) and Writing (W)**

Grammatical features		Total number of instances (S: 60 subjects W: 45 subjects)		Average number of instances per person		Average rate of occur- rence (one instance in every n words: 1/n)	
1 filler-gap construction	S	256	(incl. 27 RCs)	4.3	(0.5)	1/186	(1/1761)
	W	130	(incl. 99 RCs)	2.9	(2.2)	1/81	(1/107)
2 passive voice	S	47		0.8		1/1018	
	W	93		2.1		1/114	
3 finite coordinate clauses	S	576		9.6		1/83	
	W	177		3.9		1/60	
4 finite complement clauses	S	94	(incl. 34 that/if-clauses)	1.6	(0.6)	1/506	(1/1399)
	W	123	(incl. 64 that/whether-clauses)	2.7	(1.4)	1/86	(1/165)
5 finite adverbial clauses	S	446		7.4		1/107	
	W	152		3.4		1/70	
6 topic-comment structure	S	37		0.6		1/1285	
	W	18		0.4		1/588	
7 non-standard use of subj.-verb agreement	S	143		2.4		1/333	
	W	37		0.8		1/286	
8 non-standard use of noun number	S	180		3.0		1/264	
	W	79		1.8		1/134	
9. omission of required indefinite articles	S	120		2.0		1/396	
	W	19		0.4		1/557	

Results from Table 4.4.1 show that instances of all grammatical features except *omission of required indefinite articles* (feature 9), are more frequent in the subjects' written essays than in their speech although to varying extents. The overall results pertaining to each of the features in both the speech and written essays are addressed respectively hereinafter.

There are quite a number of instances of filler-gap constructions (feature 1; see 2.1.3.2 for details) in both the speech and written essays of the subjects. However, such filler-gap constructions in the written essays (1/81) were used over twice as frequently as in the speech (1/186). Moreover, the majority of the filler-gap constructions in the speech are very simple *wh*-questions (either direct or indirect), while the majority of such constructions in the written essays are relative clauses (99 versus 27 in the speech), which involve more complex filler-gap domains. The use of relative clauses in the written essays (1/107) is over 16 times as frequent as that in the speech (1/1761). This suggests that subjects' use of complex filler-gap constructions such as relative clauses is more characteristic of a careful/more complex style in their essay writing than of a casual/less complex style of their speech.

It is, nevertheless, significant that the use of relative clauses in both the speech and written essays conform to the Accessibility Hierarchy. In the conversation data, there are 13 instances of relativisation on SU position (48%), 10 on DO position (37%), 3 on OBL position (11%) and only one on GEN position (4%), whereas in the writing samples, there are more instances of relativisation on SU position (79; 80%) than those on DO (18; 20%) and OBL (2; 2%). The use of relative clauses in the naturalistic data therefore strongly supports the prediction made by the Accessibility Hierarchy (see 4.2.4).

The use of passive voice (feature 2)⁵⁴ in the written essays is about 9 times as frequent as in the speech (1/114 versus 1/1018), and there are more subjects who used passive voice in their written essays than in their speech (30; 67% versus 27; 45%). The incidence of passive voice also varies with respect to the range of variation, choice of words, and focus of themes in the naturalistic data. The range of the use of passive voice in the speech is between 0-5 instances, but that of the written essays is between 0-11 instances. In the speech in which the themes of conversations for all subjects remain the same, there are more instances of the use of passive voice with the same commonly used verbs such as 10 instances of *was/were born*, 5 instances of *be/get married* and 4 instances of *be used*. There are also 6 instances involving the misuse of passive voice, when active voice should be used, for instance, *I have been lived in Fushan*, and *Is the Sony's one included the remote control?*. Whereas in the written essays that vary greatly in themes of writing, the use of passive voice is characteristic of more diversities of verbs and structures. For example, unlike the 10 stereotyped instances of *was/were born* in the conversation data, two of the five instances of *to be born* in the written essays were used respectively as *No one is born a criminal* and *Also the children will be genetically born*.

The higher frequency of the use of passive voice in the written essays than in the speech seems to be reflected in the subjects' performances on the sentence combination task. In that task, among the 301 instances of relativisation from a predicted lower position to a higher position on the Accessibility Hierarchy, 35 instances (12%) involve the use of

⁵⁴ Passive voice is classified as "be + past participle" in this study, but excluded from analysis are instances such as "Train is very developed" and "It's not so crowded", in which past participles are preceded with such modifiers as *very*, *so* or *too*, and function as a predicative.

passive voice (see 4.2.1), or, put differently, the subjects used passive voice in different relativisation at the average rate of 1/95, which is quite comparable to 1/114 in the written essays. In view of the results of the subjects' use of passive voice in the sentence combination task and in their naturalistic speech and writing production, it can be concluded that passive voice is far more frequently used in L2 learners' writing than in their speech.

The use of finite coordinate clauses (feature 3), finite complement clauses (feature 4) and finite adverbial clauses (feature 5) was examined in the naturalistic data for the purpose of showing the frequencies of these clause types therein and of comparing the results with those relevant results in other tasks. Results show that all the three types of clauses occurred very frequently in both the speech and written essays except for the use of finite complement clauses in the speech (1/506), which was about 5-8 times as infrequent as the use of any other type. This shows that it is the frequency of the use of finite complement clauses rather the other two types that distinguishes the styles of the use of finite clauses between the subjects' naturalistic oral and written interlanguage production.

Concerning the filler-gap construction across a complement clause introduced by the complementisers *that*, *if* and *whether* respectively, which are comparable to the examples used for the Hierarchy for Complementisers (see (15) in 4.3.1), there is no single instance of such complex filler-gap constructions in the naturalistic data. However, the subjects did use complement clauses introduced by *that*, *if* or *whether* in their naturalistic interlanguage production, although far more frequently in their written essays (1/165) than in their speech (1/1399). Specifically, there were 29 *that*-clauses and 5 *if*-clauses in the subjects' speech, and 63 *that*-clause but only one *whether*-clause in their written essays. Even if there is no direct evidence in support of the Hierarchy for Complementisers (Hawkins 1999) due to the lack of use of filler-gap constructions across *that/if/whether*-introduced complement clauses in the naturalistic data, the use of such clauses therein still seems to lend indirect support to this hierarchy. That is, the predominance of *that*-clauses over the other two types seems to suggest that *that*-clauses are more accessible to the subjects possibly due to the less semantic content carried by

that as compared with *if* and *whether*. One point is, however, clear: a filler-gap construction across clause boundaries is indeed too complex a structure for L2 learners to master, especially in their interlanguage production.

Furthermore, the comparison between the use of finite coordinate and subordinate clauses and that of *that/if/whether*-clauses suggests that the subjects were more comfortable using the former two clause types than the latter type, which was consistent with the results in the sentence combination task (see Table 4.2.13 in 4.2.5). Among the 50 instances of various types of finite clauses produced by the subjects as alternatives to the predicted relative clauses in the sentence combination task, there are 22 finite coordinate clauses, 26 finite adverbial relative clauses, but only 2 finite complement clauses with the complementisers *that* and *if* respectively.

With respect to topic-comment structure (feature 6), although the use of such a structure in the written essays of the subjects is twice as frequent as that in their speech (1/588 versus 1/1285), the number of the subjects who use this structure in their written essays (9; 20%) is smaller than the number who use this structure in their speech (25; 42%). Among all the 37 instances of topic-comment structures and the 27 instances of relative clauses in the subjects' speech, there are only two examples of relative clauses that relativise on a noun in the topic component of a topic-comment structure. Similarly, there is only one such example among all the 18 instances of topic-comment structures and the 99 instances of relative clauses in the subjects' written essays. Similar results can also be found in the repetition and sentence combination tasks: only seven subjects (out of 116 subjects in the two tasks) produced 17 instances (out of 1504 instances in the two tasks) of relativisation on a noun in the topic component of a topic-comment structure (see Table 4.2.15 in 4.2.5.1). Overall, the use of topic-comment structure is a fairly common phenomenon amongst the subjects in their naturalistic interlanguage production; however, when the topic component involved relativisation, such a structure became so structurally complex in terms of an extra filler-gap construction for processing that only a few subjects used it.

I have included an analysis of non-standard use of subject-verb agreement (feature 7), non-standard use of noun number (feature 8) and omission of required indefinite articles (feature 9) because these three types of non-standard forms, among others, are quite characteristic of both the spoken and written interlanguage of the subjects. The three grammatical features (non-standard forms) can therefore be used as indices of the grammatical ability of the subjects and can be compared with the results of their use of ungrammatical non-standard forms in other tasks so as to gain a comprehensive picture of the grammatical features of their interlanguage.

In the conversation data, the length of speech of the subjects ranges between 154 to 1742 words with an average of about 800 words per person, and examples exemplifying one or more of the three types of non-standard forms are found in the speech of every subject (see **Appendix X**). Whereas in the written essays, the length of essays of the subjects ranges between 55 to 498 words with an average of 235 words per person, and there are seven subjects who did not produce any instance of the three types of non-standard forms in their writing (see **Appendix XI**). The results (Table 4.4.1) showed that the non-standard use of subject-verb agreement and noun number in the written essays (1/286 and 1/134 respectively) was a little more frequent than that in the conversation data (1/333 and 1/264 respectively), but the omission of required indefinite articles in the written essays was a little more infrequent than that in the conversation data (1/557 versus 1/396). However, there were bigger differences between individual performances on these three types, for example, the non-standard use of noun number ranged between 0-11 instances in the conversation data and between 0-9 instances in the written essays (see **Appendices X and XI**). Despite these individual differences, the non-standard use of subject-verb agreement and of noun number, and the omission of required indefinite articles are common interlanguage features of the subjects in both their speech and their written essays.

4.4.2. Results Examined by Different English Proficiency Level Groups

Now we examine the extent to which different English proficiency level groups may differ from one another with regard to the use of grammatical features examined in 4.4.1. Results of the examination are provided in Table 4.4.2.

Table 4.4.2 Counts and Rates for the Use of Different Grammatical Features in Naturalistic Speech (S) and Writing (W) by Different English Proficiency Level Groups

Grammatical features		English Proficiency Levels					
		Low (S: 22 subjects W: 16 subjects)		Mid (S: 19 subjects W: 16 subjects)		High (S: 19 subjects W: 13 subjects)	
1 filler-gap construction	S	51	(1/307)*	96	(1/168)	107	(1/148)
	W	32	(1/102)	35	(1/106)	63	(1/57)
2 passive voice	S	13	(1/1203)	17	(1/946)	17	(1/932)
	W	17	(1/193)	29	(1/128)	47	(1/76)
3 finite coordinate clauses	S	186	(1/84)	184	(1/87)	206	(1/77)
	W	52	(1/63)	68	(1/55)	57	(1/63)
4 finite complement clauses	S	24	(1/651)	32	(1/503)	38	(1/417)
	W	43	(1/76)	36	(1/103)	44	(1/82)
5 finite adverbial clauses	S	135	(1/116)	180	(1/89)	131	(1/121)
	W	54	(1/61)	47	(1/79)	51	(1/70)
6 topic-comment structure	S	16	(1/977)	13	(1/1237)	13	(1/1218)
	W	2	(1/1638)	8	(1/464)	8	(1/448)
7 non-standard use of subject-verb agreement	S	46	(1/340)	47	(1/342)	50	(1/317)
	W	8	(1/409)	14	(1/265)	15	(1/239)
8 non-standard use of noun number	S	64	(1/244)	57	(1/282)	59	(1/268)
	W	24	(1/136)	30	(1/124)	25	(1/143)
9. omission of required indefinite articles	S	43	(1/364)	42	(1/383)	35	(1/453)
	W	5	(1/655)	11	(1/371)	4	(1/897)

* The left figure indicates the total number of instances of a grammatical feature used by a group, while the right figure (within parentheses) refers to the average rate of the use of that grammatical feature in terms of "once in every *n* words", hence comparable across groups.

With respect to filler-gap constructions (feature 1), the high-level group used such constructions far more frequently than the low-level group in their speech and significantly more frequently than both the low-level and mid-level groups in their written essays. The mid-level group used such constructions significantly more frequently than the low-level group in their speech. This suggests that the more proficient in English the subjects are, the more frequently they use filler-gap constructions in their naturalistic speech and writing.

With the use of passive voice (feature 2), again, the high-level group used more instances of passive voice than the other two groups in both their speech and written essays. In the

written essays in particular, the high-level group used passive voice more frequently than the mid-level group, which, in turn, used passive voice more frequently than the low-level group. A close examination, however, shows that the apparent differences in the use of passive voice in the written essays between different level groups can be attributed more to the differences of the types and themes of the subjects' written essays than to the level differences. For example, there was only one instance of the use of passive voice found in the written essays of seven low-level subjects with a frequency (1/739) far lower than this group's average frequency (1/193); however, all of the seven essays were personalised narratives with the same topic "Who am I?". Two subjects (B1 and V5) from the high-level group, on the other hand, used passive voice more frequently (10 and 11 instances respectively) than any other subjects in the study with a combined frequency (1/19) far higher than this group's average frequency (1/76); their written essays, entitled "Benefits given to man-kind through space exploration" and "Explaining a diagram" respectively, were more of a depersonalised nature involving argumentation and description respectively. In the conversation task in which the themes of the conversation remained the same for all the subjects, the differences between different level groups in the use of passive voice are not as big.

There is no big difference between the three level groups in the use of finite coordinate, subordinate and complement clauses (features 3, 4 and 5) in either the conversation data or the written essays. In fact, each of the three groups used two of the six features (features 3, 4 and 5 in speech and writing respectively) marginally more frequently than the other two groups (see Table 4.4.2). This suggests that the L2 learners of English in this study performed similarly in the use of varieties of finite clauses regardless of their English proficiency levels.

Again, there is no big difference between the three level groups in the use of topic-comment structures (feature 6) in either the conversation data or the written essays. However, the low-level group used such structures marginally more frequently in their speech and less frequently in their written essays than the other two groups (see Table 4.4.2).

With regard to the non-standard use of subject-verb agreement (feature 7), noun number (feature 8) and indefinite articles (feature 9), it is a bit surprising that there are no big differences between different level groups in either the conversation data or the written essays. However, amongst the six features (features 7, 8 and 9 in speech and writing respectively), the high-level group used non-standard forms of noun number in written essays and non-standard forms of indefinite articles in both speech and written essays marginally less frequently than the other two groups; the mid-level group used non-standard forms of subject-verb agreement in speech and non-standard forms of noun number in speech marginally less frequently than the other two groups; and the low-level group used non-standard forms of subject-verb agreement in writing marginally less frequently than the other two groups (see Table 4.4.2). The overall results therefore seem to be indicative of the proficiency levels of English grammar for the three different level groups, albeit weakly. That is, overall, the high-level group performed marginally better in the grammatical aspects under examination than the mid-level group, which, in turn, performed marginally better than the low-level group.

4.4.3. Results Examined by Different First Language Groups

We now examine the extent to which different L1 groups may differ from one another with regard to the grammatical features examined in 4.4.1. Results of the examination of the conversation data are provided in Table 4.4.3.

Table 4.4.3 Counts and Rates for the Use of Different Grammatical Features in the Conversation Data by Different First Language Groups

Grammatical features	First Languages*							
	M (N = 18; 836)	J (N = 11; 651)	K (N = 8; 801)	V (N = 7; 924)	T (N = 7; 698)	In (N = 3; 777)	R (N = 5; 919)	B (N = 1; 652)
1 filler-gap construction	71** 1/212	36 1/199	38 1/169	26 1/249	28 1/174	15 1/155	28 1/164	14
2 passive voice	20 1/753	2 1/3582	11 1/583	5 1/1293	3 1/1628		3 1/1531	3
3 finite coordinate clauses	166 1/91	121 1/59	86 1/75	64 1/101	54 1/90	19 1/123	60 1/77	6
4 finite complement clauses	27 1/558	11 1/651	6 1/1068	17 1/380	5 1/977	9 1/259	16 1/287	3
5 finite adverbial clauses	151 1/100	70 1/102	48 1/134	68 1/95	52 1/94	20 1/117	33 1/139	4
6 topic-comment structure	14 1/1075	8 1/895	3 1/2136	7 1/928	3 1/1628		2 1/2297	
7 non-standard use of subj.-verb agreement	34 1/443	14 1/512	21 1/305	37 1/175	17 1/287	4 1/583	15 1/306	1
8 non-standard use of noun number	52 1/290	24 1/298	18 1/356	27 1/239	26 1/188	8 1/291	18 1/255	7
9. omission of required indefinite articles	28 1/538	27 1/265	22 1/291	19 1/340	9 1/543	3 1/777	11 1/418	

* First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, B = Bangla, and R = Romance including Spanish, French and Italian. The figures within parentheses refer respectively to the number of participants and the average number of words spoken per participant in each group.

** The upper figure indicates the total number of instances of a grammatical feature used by a group, while the lower figure refers to the average rate of the use of that grammatical feature in terms of "once in every *n* words", hence comparable across groups.

Results show that, generally, the interlanguage features under investigation are characteristic of all L1 groups, albeit to varying degrees. Specifically, the Mandarin-speaking group performed around average in all categories. The Japanese-speaking group had the highest frequency of finite coordinate clauses (1/59), topic-comment structures (1/895) and also the omission of required indefinite articles (1/265). The Korean-speaking group had the highest frequency of utterances containing passive voice (1/583), but the lowest frequency of finite complement clauses (1/1068) and non-standard forms involving noun number (1/356). The Vietnamese-speaking group had the highest frequency of non-standard forms in subject-verb agreement (1/175) and one of the two highest frequencies of finite adverbial clauses (1/95), but the lowest frequency of filler-gap constructions (1/249). The Thai-speaking group had the highest frequency of finite adverbial clauses (1/94), and the highest frequency of non-standard forms concerning noun number (1/188). The Indonesian-speaking group did not produce any utterances

using passive voice and topic-comment structure, they had the highest frequency of filler-gap constructions (1/155) and finite complement clauses (1/259), but the lowest frequency of finite coordinate clauses (1/123), non-standard forms of subject-verb agreement (1/583) and omission of required indefinite articles (1/777). Finally, the Romance-speaking group had the second highest frequency of finite complement clauses (1/287) but the lowest frequency of finite adverbial clauses (1/139).

We now turn to the examination of these grammatical features in the written essays by different L1 groups, and the results are provided in Table 4.4.4.

Table 4.4.4 Counts and Rates for the Use of Different Grammatical Features in the Written Essays by Different First Language Groups

Grammatical features	First Languages*							
	M (N = 15; 273)	J (N = 6; 152)	K (N = 5; 175)	V (N = 5; 303)	T (N = 5; 246)	In (N = 3; 271)	R (N = 4; 180)	Others (N = 2; 206)
1 filler-gap construction	50** 1/82	12 1/76	6 1/146	21 1/72	17 1/72	16 1/51	3 1/240	5
2 passive voice	24 1/171	8 1/114	13 1/67	16 1/95	11 1/112	5 1/162	6 1/120	10
3 finite coordinate clauses	69 1/59	14 1/65	18 1/49	26 1/58	13 1/95	15 1/54	16 1/45	5
4 finite complement clauses	51 1/80	15 1/61	9 1/97	16 1/95	13 1/95	7 1/116	9 1/80	3
5 finite adverbial clauses	64 1/64	16 1/57	12 1/73	13 1/116	16 1/77	12 1/68	11 1/65	8
6 topic-comment structure	4 1/1026	2 1/457		4 1/378		4 1/203	4 1/180	
7 non-standard use of subj.-verb agreement	20 1/205	3 1/304	6 1/146	3 1/504	4 1/308	1 1/812		
8 non-standard use of noun number	28 1/147	9 1/101	13 1/67	5 1/303	18 1/68	1 1/812	4 1/180	1
9 omission of required indefinite articles	8 1/513	1 1/913	2 1/437	1 1/1513	3 1/411	2 1/406	1 1/719	1

* First languages: M = Mandarin, J = Japanese, K = Korean, V = Vietnamese, T = Thai, In = Indonesian, R = Romance including Spanish, French and Italian, and *Others* (listed but not for analysis) include two subjects speaking Bangla and Malay respectively with the Malay-speaking subject participating only in essay collection. The figures within parentheses refer respectively to the number of participants and the average number of words written per participant in each group.

** The upper figure indicates the total number of instances of a grammatical feature used by a group, while the lower figure refers to the average rate of the use of that grammatical feature in terms of "once in every *n* words", hence comparable across groups.

Results from Table 4.4.4 show that, again, the interlanguage features examined in the written essays of the subjects are characteristic of all L1 groups though the absolute scores in terms of occurrence rates for different grammatical features vary from group to

group. Similar to the results from the conversation data, the Mandarin-speaking group performed around average in all features except passive voice in the written essays; they had the lowest frequency of the use of passive voice (1/171). Generally, this group performed consistently between the highest and the lowest scores in all types of tasks. For example, they had the second highest score of *(near) correct repetition* (31.5%) in the repetition task (see Table 4.2.3 in 4.2.1), the fourth highest score of *predicted RC embedding* (45%) in the sentence combination task (see Table 4.2.6 in 4.2.2) and the third highest score of correct responses to RC-related items (68%) in the grammaticality judgment task (see Table 4.2.10 in 4.2.3). However, in terms of making ungrammatical non-standard forms in the repetition and sentence combination tasks, the Mandarin-speaking group scored the best in the former (5.3 forms per person) and the second best in the latter (2.8 forms per person) (see 4.2.5.3).

The Japanese-speaking group, in contrast to their performances in the conversation (having the highest frequency of finite coordinate clauses and omission of required indefinite articles), had the highest frequency of finite complement clauses (1/61) and finite adverbial clauses (1/57) but one of the lowest in omission of required indefinite articles (1/913) in the written essays. It shows that the Japanese-speaking subjects tend to use a lot more finite coordinate clauses in their speech and a lot more finite complement clauses and adverbial clauses in their written essays as compared with the subjects of the other L1 groups. It also shows that they omit required indefinite articles quite frequently in their speech, but when they write in English, they rarely omit such articles.

Considering their performances in other tasks, it is not surprising that the Japanese-speaking subjects should behave like this. They had the second lowest score of *(near) correct repetition* (22%) in the repetition task, but the highest score of *predicted RC embedding* (56%) in the sentence combination task and the second highest score of correct responses to RC-related items (70%) in the grammaticality judgment task. Similarly, the Japanese-speaking group had the second highest rate of making ungrammatical non-standard forms in the repetition task (7.9 forms per person) but the third lowest in the sentence combination task (2.9 per person). On the whole, the

Japanese-speaking subjects performed consistently across all the tasks: in the repetition and conversation (involving listening and speaking), they made more ungrammatical non-standard forms; but in the sentence combination, grammaticality judgment and writing tasks (involving writing and grammatical knowledge), they made much fewer such forms.

The Korean-speaking group had the highest frequency of the use of passive voice (1/67), the non-standard use of subject-verb agreement (1/146), the non-standard use of noun number (1/67), and one of the highest frequencies of the omission of required indefinite articles (1/437) (see Table 4.4.4). They consistently used passive voice in their speech and written essays more frequently than any other group; however, surprisingly, their use of non-standard forms of noun number was more frequently than the other groups in the written essays but less frequently than the other groups in the conversation data.

It is even more surprising if we look at their performances in the other three tasks, in which they performed exactly the same as the Japanese-speaking groups. They had the lowest score of *(near) correct repetition* (19%) in the repetition task, but the second highest score of *predicted RC embedding* (54%) in the sentence combination task and the highest score of correct responses to RC-related items (71%) in the grammaticality judgment task. Generally, the Korean-speaking subjects performed much better in the sentence combination and grammaticality judgment tasks than in the repetition task; in naturalistic production of English, they, nonetheless, performed better in speech than in writing in the use of noun number – one of the indices for measuring the subjects' English grammatical knowledge. The examination of the performances of the Korean-speaking group across different tasks shows that the interlanguage features of this group are inconsistently variable.

As for the Vietnamese-speaking group, two features stand out in comparing their performances in speech and writing. The first has to do with the use of finite adverbial clauses: they produced such clauses most frequently in their speech (1/95) but least frequently in their writing (1/116) as compared with other L1 groups. The second

involves use of non-standard forms in speech and writing. In writing, they had the lowest frequency in omission of required indefinite articles (1/1513), the second lowest in non-standard use of noun number (1/303) and the third lowest in non-standard use of subject-verb agreement (1/504) (see Table 4.4.4). In speech, on the other hand, the mirror image occurs: they had the highest frequency in non-standard use of subject-verb agreement (1/175), the second highest in non-standard use of noun number (1/239) and the third highest in omission of required definite articles (1/340) (see Table 4.4.3). This suggests that the Vietnamese-speaking subjects used non-standard forms less frequently in their written essays but more frequently in their speech than the other groups.

Although the Vietnamese-speaking subjects performed consistently at the average level across the tasks of repetition, sentence combination and grammaticality judgment in terms of predicted correct responses, their use of ungrammatical non-standard forms in the repetition, sentence combination and grammaticality judgment tasks shows a different picture. While they had an average of ungrammatical non-standard forms (7 per person) in the repetition task, they had a surprisingly low rate of such forms in the sentence combination task – 0.7 forms per person (the lowest), which is far better not only than the highest rate (5.3) but than the second lowest rate (2.8) as well. In judging RC-related non-standard forms in the grammaticality judgment task (see Table 4.2.17), they had the second lowest percentage (13%) of judging those forms as acceptable – the second best after the Indonesian-speaking group. All this seems to indicate that the Vietnamese-speaking subjects generally performed well in grammatical aspects of English across different tasks particularly in sentence combination, grammaticality judgment and written essays; however, when they produced spontaneous speech, ungrammatical non-standard forms accrue drastically at least in the aspects under examination. Their better performances in grammatical forms has been suggested as a function of several factors including their unique experience of learning English in their home country (see 4.1.4, 4.2.2, and 4.2.5.3 for details).

The Thai-speaking group had the lowest frequency of finite coordinate clauses (1/95) in their written essays in comparison with the highest frequency of finite adverbial clauses

(1/94) in their speech. Prominently, they had one of the two highest frequencies in both non-standard use of noun number (1/68) and omission of required indefinite articles (1/411) in their written essays, which is consistent with the highest frequency in non-standard use of noun number (1/188) in their speech. It shows that this group tends to use finite adverbial clauses more frequently in speech but finite coordinate clauses more frequently in written essays, yet in both their speech and written essays, they seem to have difficulties using noun number correctly more frequently as compared with other L1 groups.

In the other three tasks, they achieved the best score of *(near) correct repetition* (32.1%) in the repetition task, the fifth in *predicted RC embedding* (33%) in the sentence combination task and the joint fourth score of correct responses to RC-related items (63%) in the grammaticality judgment task. Concerning ungrammatical non-standard forms made by different L1 groups in terms of ranking from low to high rates, the Thai-speaking group, however, comes out the third (6.4 forms per person) in the repetition task, and the sixth (4.4 forms per person) in the sentence combination task. It therefore suggests that no matter whether the Thai-speaking subjects scored best as in *(near) correct repetition* in the repetition task or around the average as in other tasks, they were prone to using ungrammatical non-standard forms a bit more frequently than expected.

With regard to the Indonesian-speaking group, their performances in written essays as well as speech exhibit the following features. First, they tended to use filler-gap constructions most frequently in both their speech (1/155) and their written essays (1/51). Second, they had the second lowest frequency of the use of passive voice in their written essays (1/162) and the lowest in their speech; actually they did not produce any passive utterances in the conversation. Third, they used finite complement clauses most frequently (1/259) in their speech, but least frequently (1/116) in their writing. Finally, both their speech and writing are characteristic of very few non-standard forms (19 instances in total): in writing, they had the lowest frequency in non-standard use of noun number (1/812), the second lowest in non-standard use of subject-verb agreement (1/812) but the highest in omission of required indefinite articles (1/406: but only 2 tokens in

total); in speech, they had the lowest frequency in subject-verb agreement (1/583) and omission of required indefinite articles (1/777).

In the other three tasks, the performances of the Indonesian-speaking group varied from task to task. In the repetition task, they were among the top three groups in terms of *(near) correct repetition* (30% with the other two being 31.5% and 32.1% respectively; see Table 4.2.3 for details). However, they came out penultimate (the sixth) both in the sentence combination task in terms of *predicted RC embedding* (26%) and in the grammaticality judgment task (related to types of relative clauses) in terms of correct judgment (47%). Notwithstanding in terms of ranking the rates/percentages of making ungrammatical non-standard forms from low to high, they came out the second (6 forms per person) in the repetition task, the fifth (3.7 forms per person) in the sentence combination task, and the first (7%) in the grammaticality judgment task (related to types of non-standard forms of relative clauses). All in all, the Indonesian-speaking subjects seem to be more competent in naturalistic production of English (speech and writing) and the repetition task than in the sentence combination and grammaticality judgment tasks (related to types of relative clauses) in terms of predicted results; nevertheless, they are even more competent in avoiding ungrammatical non-standard forms across all tasks.

Finally, the Romance-speaking group had the highest frequency of finite coordinate clauses (1/45) but the lowest frequency of filler-gap constructions (1/240) and of non-standard use of subject-verb agreement (in fact, none at all) in their writing. They also had the second highest frequency in finite complement clauses both in their writing (1/80) and in their speech (1/287). Apart from the above results, most statistics regarding the categories examined for this group in both speech and writing, fall at the average level, i.e. either at the third or fourth place (among 7 L1 groups) ranked in terms of occurrence frequencies of a category. For example, regarding the frequencies of non-standard forms (categories 7 and 8) made by this group in both speech and writing, all the frequencies are ranked either as third or fourth except the frequency for non-standard use of subject-verb agreement in the writing, which is ranked the first – the lowest. All this

suggests that the Romance-speaking group performed at least at the average level as compared with all the other groups in both speech and writing.

The performances of the Romance-speaking group in speech and writing are, however, in sharp contrast to their performances in the other three tasks. In terms of predicted correct responses in the three tasks, they came out equal fifth (22%) in the repetition task, and last in both sentence combination (13%) and grammaticality judgment tasks (related to types of relative clauses; 45%). In terms of ungrammatical non-standard forms made in the three tasks – repetition, sentence combination, and grammaticality judgment (related to types of non-standard relative clauses), they fared even worse: they scored the highest rate/percentage in all the three tasks. Tentative explanation for all this has been suggested in section two of this chapter (see 4.2.5.3 for details).

The striking differences between their performance on the naturalistic speech and writing on the one hand, and those on the other three tasks on the other, can be accounted for partly by intragroup differences and partly by the nature of the tasks in point. For example, the intragroup differences between the three Spanish-speaking subjects and the other two (French and Italian) in the conversation are greater than similar differences in any other task. The three Spanish-speaking subjects produced 43 instances of the non-standard use of subject-verb agreement and noun number, and the omission of required indefinite articles in their speech, whereas the two French- and Italian-speaking subjects produced only 3 such instances in their speech (see **Appendix X**). So is the case in the writing in which the three Spanish-speaking subjects produced five such instances but the Italian-speaking subject produced none (see **Appendix XI**). Furthermore, the groups as a whole seems to perform better in naturalistic data than in the other three tasks; even the Spanish-speaking subjects seem to improve their performances in writing.

4.4.4. Summary

Interlanguage features in relation to some of the findings in the previous sections (4.1-4.3) are examined in the naturalistic data – the speech and written essays of the subjects in this study. Results of the examination are summarised below.

Filler-gap constructions are widely used in L2 learners' speech and written essays, and those used in writing are far more complex than those used in speech. Although relative clauses are the most complex filler-gap constructions produced by the subjects, they are mostly relativised on the higher positions – SU and DO – on the Accessibility Hierarchy with only five (out of 115 instances) relativised on OBL position. There are no exemplifying examples of relativisation on the lower positions – GEN and OCOMP – on the hierarchy in the subjects' speech and writing, let alone those more complex filler-gap constructions that are tested in the grammaticality judgment task such as those involving two gaps.

Passive voice, a means via which many subjects relativised on a position higher on the Accessibility Hierarchy than the one predicted in the sentence combination task, is found to be common in the written essays of the subjects rather than in their speech. In effect, the use of passive voice in instances of *different relativisation* in the *sentence combination* task but not in the *repetition* task, can be at least partly ascribed to the subjects' preference of using passive voice in their writing rather than in their speech.

Finite coordinate and adverbial clauses, which the subjects used most frequently when they failed to produce the required relative clauses in the sentence combination task (see Table 4.2.13), are, in fact, used very frequently in both their speech and written essays. The finite complement clauses, on the other hand, are used far more frequently in the subjects' written essays than in their speech. With respect to finite complement clauses introduced by the complementisers *that*, *if*, and *whether*, there is a predominant preference of *that*-clauses (92 instances) over *if*- (5 instances) and *whether*-clauses (1 instance only), which lends indirectly to the Hierarchy for Complementisers (Hawkins 1999).

Although the use of topic-comment structure in the repetition and sentence combination tasks is confined to only a couple of Korean-, Vietnamese- and Thai-speaking subjects (see Table 4.2.17), this structure is widely available to all L1 groups in their naturalistic

English interlanguage production (i.e. in speech and/or written essays). Relativisation on a noun in the topic (comparable to those in the repetition and sentence combination tasks) is, however, very infrequent both in the speech (2 out of 37 instances) and in the written essays (1 out of 18 instances) of the subjects. Therefore all the results concerning the use of topic-comment structure in the study suggest that topic-comment structure is a common interlanguage feature among different L1 groups including the Romance-speaking subjects in their naturalistic speech and writing, yet that when the topic involves complex filler-gap constructions such as relative clauses, such a structure is found far less frequently both in the repetition and sentence combination tasks and in their speech and written essays.

The three features of non-standard use of subject-verb agreement, noun number and indefinite articles were intended to be used as indices for confirming differences between the English language proficiency of the subjects placed in different instructional levels. However these features were found to be very similar in both the speech and written essays of the subjects (though with varying degrees). However, the subjects' use of such non-standard forms in naturalistic data is only weakly indicative of the level of English grammar of different level groups, and is generally consistent with the use of ungrammatical non-standard forms by different L1 groups in other tasks.

In closing, the examination and analysis of the naturalistic speech and written essays of the subjects with respect to their overall performances, English language proficiency levels and first languages, help to address some of the issues that emerged from the deductively-contrived tasks of this study, thereby providing a richer picture of the interlanguage features under investigation.

Chapter Five Summary of Major Theoretical Findings

This chapter summarises the theoretical findings of the research, which focuses on the study of syntactic features of the English interlanguage of ESL learners. Section one (5.1) provides a summary of the findings concerning the acquisition of English interrogative questions, section two (5.2) of those concerning the acquisition of English restrictive relative clauses, section three (5.3) of those concerning the acquisition of various filler-gap constructions, and section four (5.4) summarises the most significant findings of this study.

5.1. Findings regarding the Acquisition of English Interrogative Questions

Comparable to a study conducted by Eckman et al (1989), which tested Greenberg's (1963) implicational universal concerning questions via an elicitation task among L2 learners of English, this study further examines the same universal in a naturalistic task among interlanguage learners of English with more diverse L1-speaking backgrounds. This study further examines whether L2 learners' English proficiency levels and first languages may have some impact on their performances relating to the universal under examination. Greenberg's implicational universal regarding constituent order of *yes/no* and *wh* questions, which was postulated on the basis of a 30-primary-language sample and was first testified in interlanguage data by Eckman et al (1989), was again validated by the interlanguage data in this study. That is, the universal "*yes/no* inversion implies *wh* inversion which, in turn, implies initial *wh* question word" is predominantly supported by the performance of the subjects in this study regardless of their English proficiency levels and their respective first languages.

Both this study and the study by Eckman et al (1989) have one exception to the "*yes/no* inversion implies *wh* inversion", yet it is significant that the overall results of both studies overwhelmingly support the universal. It is even more significant that despite the differences in the performances between different proficiency groups (low, mid and high)

and between different L1 groups (7 in all), the performance of all these groups supports the universal. The fact that the three different proficiency groups performed all in the direction of structural conformity to the universal seems to suggest that interlanguage developmental factors do not affect the validity of the universal under test. Similarly, the factor of first language transfer does not seem to affect the performances of all L1 groups, whether the implicans and/or implicatum of the universal are exhibited systematically in their respective first languages or not.

All this seems to suggest, however, that it is the *Interlanguage Structural Conformity Hypothesis* – all universals that are true for primary languages are also true for interlanguages (Eckman et al 1989: 195) – that is accountable for the results of the study. In other words, all groups conform to the implicational universal relating to interrogative questions no matter how high or low percentages each group scores, whether they are at the low, mid or high proficiency levels, and whether their first languages similar to or different from English. In sum, the findings relating to the implicational universal lend strongly to the claim that implicational universals that hold true for first languages should also hold true for second languages, and also to the implication that interlanguages, like primary languages, are processed alike by human beings.

5.2. Findings regarding the Acquisition of English Restrictive Relative Clauses

Studies on the acquisition of English restrictive relative clauses by L2 learners with respect to the Accessibility Hierarchy have been well documented in the literature of SLA (see 2.2.2 for details). A large part of the present study is devoted to the research in this area. What differentiates this study from the previous ones is its focus on a processing-motivated typological approach and its methodology of cross-task/style data collection and analysis. The processing approach proposed by Hawkins (1994, 1999) which associates processing load with structural complexity in filler-gap constructions is adopted in addressing the data. The data under examination include those collected from the naturalistic production of speech and writing of the subjects as well as from the

analytically and deductively designed tasks – elicited repetition, sentence combination and grammaticality judgment.

Results of the study show that, overall, the Accessibility Hierarchy is supported in all types of data, albeit to varying degrees, that is, predicted instances of relativisation on SU position are more than those on DO position, which are, in turn, more than those on IO/OBL positions, and so on, with the exception of GEN position in the latter three tasks. The Accessibility Hierarchy is supported in terms of *(near) correct repetition* in the repetition task (see 4.2.1), in terms of *predicted RC embedding* in the sentence combination task (see 4.2.2), in terms of correct instances of judgment in the grammaticality judgment task (see 4.2.3), and in terms of use of RCs in the subjects' speech and writing (see 4.2.4). Support for the Accessibility Hierarchy is also found within each English proficiency level and within each first language group, for which there was relatively sufficient data. Even when the subjects provided relativisations that were different from the predicted ones their performances conformed to the Accessibility Hierarchy – among the 426 instances of alternative relativisations in the repetition and sentence combination tasks, 414 involve relativisation on a higher position with only eleven on a similar position and one on a lower position.

As for the subjects' performances relating to relativisation on GEN position in different tasks, the results show that GEN position is treated as more accessible to relativisation than DO, IO/OBL and OCOMP positions in the repetition and sentence combination tasks, and even more so than all the other positions including SU in the grammaticality judgment task. Nevertheless, this exception to the Accessibility Hierarchy, presumably caused by the unique characteristic of the English genitive relative marker *whose* (Gass 1979), does receive a more natural explanation when viewed in terms of the processing load associated with the structural complexity of this position as well as its cognisant salience. In fact, similar cases have been well documented elsewhere in the literature of SLA studies (e.g. Gass 1979; Li and Li 1994). In contrast, however, this high level of accessibility evident in the three deductively designed tasks is not able to be assessed in

the naturalistic speech and writing of the subjects because there was only one instance of relativisation on GEN position which involved the use of an incorrect RC marker.

In view of all the results and analyses, we can conclude that the Accessibility Hierarchy is overwhelmingly supported in the repetition and sentence combination tasks in terms of both the predicted RC formation and different relativisation regardless of English proficiency levels and first languages of the subjects, and that it is also supported in the grammaticality judgment task and their speech and writing as well. Thus, the research shows that the Accessibility Hierarchy is a valid predictor for L2 performances.

It is significant that the processing explanation for the Accessibility Hierarchy (Hawkins 1999) provides a more unifying approach than others in addressing types of regularities and irregularities found in the data under examination. For example, the subjects' performances in conformity with the Accessibility Hierarchy in various tasks are well accounted for in terms of processing load associated with the structural complexity of each relativised position. While *the Absolutive Hypothesis* (Fox 1987) seems to account better for the use of relative clauses in the subjects' speech, it fails to address the same phenomenon in their writing; the Accessibility Hierarchy, on the other hand, can validly predict both cases. Furthermore, the processing-motivated explanation can adequately and convincingly address the case of irregularity pertaining to the sentence combination task (see 4.2.2), for which the *Interruption* account (Sadighi 1994) fails to find an answer even though it can address adequately many cases in that task. In short, the Accessibility Hierarchy interpreted from a processing perspective provides a unifying and natural explanation for predicting and interpreting interlanguage data with respect to relative clauses in this study.

5.3. Findings regarding Other Implicational Universals

Six of the implicational hierarchies/hypotheses, either revisited or newly motivated by Hawkins (1999), are tested for the first time against interlanguage data in the grammaticality judgment task of the study (see **Appendix 1**). The subjects' performances

pertaining to these typological universals are presented in **Appendix IX**, and detailed analyses of their performances are provided in **4.3**.

The validity of these implicational universals as predictors for interlanguage performances is justified in most cases in this study. While the subjects' performances fully support the Subordinate Gap/No Gap hierarchy, the Hierarchy for Complementisers, and the Bridging Verb Hierarchy, and mostly support the Clause Embedding Hierarchy and the Valency Completeness Hypothesis, they disconfirm the Head Noun Phrase Hierarchy.

The evidence from the grammaticality judgment task in the present study fully supports the Subordinate Gap/No Gap hierarchy, the Hierarchy for Complementisers, and the Bridging Verb Hierarchy. With regard to the Hierarchy for Complementisers, the filler-gap construction across a *that*-clause is regarded as more acceptable (albeit slightly) than that across an *if*-clause, which is, in turn, regarded as significantly more acceptable than a *whether*-clause (see **4.3.1**). There is also indirect evidence in support of the hierarchy in the naturalistic data. In spite of the non-existence of a filler-gap construction across a *that/if/whether*-clause in the naturalistic speech and writing, the subjects' use of the three complementisers in sentential complements (see **4.4.1**) complies (albeit to varying extents) with their performances in the grammaticality judgment task. As regards the Bridging Verb Hierarchy, this study found that the filler-gap construction across a *that*-clause with the matrix verb *say* is regarded as significantly more acceptable than filler-gap constructions with either the matrix verb *say softly* or the matrix verb *whisper*, both of which are more semantically loaded than *say* and are judged as equally acceptable, hence fully validating the hierarchy (see **4.3.3**). With respect to the Subordinate Gap/No Gap hierarchy, the filler-gap construction across an NP complex with one gap within its filler-gap domain, is regarded as significantly more acceptable than that with two gaps within its filler-gap domain (see **4.3.5**). In fact both constructions are ungrammatical in English; therefore the subjects' performances in this regard are even more significant in that they strongly support not only the validity of the hierarchy but its underlying processing motivation as well.

In the case of the Clause Embedding Hierarchy and Valency Completeness hypothesis, this study provides two sets of confirming data and one set of undiscriminating data for each of these implicational universals. Interestingly, however, the structure of filler-gap constructions in the set of undiscriminating data for both universals is far more complex than the structure in the confirming sets of data. For the Clause Embedding Hierarchy, the filler-gap construction of *wh*-islands with two gaps in the complementiser phrases (in the discriminate set), is structurally more complex than the filler-gap construction with one gap in an infinitival phrase, finite clause or complex NP (in the confirming sets) (see 4.3.4). For the Valency Completeness Hypothesis, the filler-gap domain across a sentential complement/subject (in the indiscriminate set) is structurally more complex than the filler-gap domain across a prepositional complement or infinitival complement (in the confirming sets) (see 4.3.6). All this suggests that these two universals do hold for the interlanguage data unless the filler-gap constructions such as *wh*-islands and sentential subjects are too complex for L2 learners to comprehend.

The Head Noun Phrase Hierarchy was not found to hold for the interlanguage data under examination. According to the hierarchy, the filler-gap construction across an NP with an indefinite head *a* (less semantically specific) should be regarded as more acceptable than that with a definite head *the*, which should, in turn, be regarded as more acceptable than that with a possessive modifier such as *John's*. However, contrary to these predictions, the subjects in this study judged the filler-gap construction with *a* as the most unacceptable despite the fact that they treat the filler-gap construction with *the* as more acceptable than that with *John's* (see 4.3.2).

To sum up, the implicational hierarchies and hypotheses of Hawkins (1999) tested in this study are predominantly supported by the interlanguage data in this study. All this evidence suggests that processing-motivated implicational universals/hierarchies of Hawkins (1999), not only provide a rich source of hypothetical assumptions to be tested against L2 data, but prove to be valid predictors for L2 acquisition as well. On the other hand, the theoretical implications drawn from SLA studies in this regard have to do with

the applicability of these universals in secondary languages such as interlanguage, hence lending support to the universality of the universals. That is, if the universals are upheld in L2 studies, they are all the more predictive.

5.4. Summary

Significantly, this study, which focused on the examination of syntactic features of the English interlanguage of ESL learners via typological (implicational) universals, overwhelmingly confirms the validity and predictive power of these universals for interlanguage phenomena. The typological (implicational) universals under examination include the word order universal pertaining to question formation (Greenberg 1963), the Accessibility Hierarchy concerning RC formation (Keenan and Comrie 1977), and a number of recently-formulated implicational hierarchies/hypotheses (Hawkins 1999), which are predominantly supported by the results of the study. Variable as the performances of the subjects in this study may be in terms of English proficiency levels, first languages, and/or types (styles) of data, these universals are still found to be, on the most part, valid predictors for interlanguage behaviour.

It is no less significant that the processing account in terms of filler-gap domains (Hawkins 1999) can not only explain consistently why typological universals are postulated the way they are, but can adequately address both regularities and irregularities of interlanguage performances relating to the universals as well. As compared with other approaches, this processing approach promises a more unifying, consistent and natural account for the interlanguage data at least as has been demonstrated in this study, which suggests that universal processing constraints may play a substantial role in second language acquisition as well in first language acquisition.

Chapter Six Concluding Remarks

This chapter provides some concluding remarks concerning the findings of this study. Section one (6.1) addresses some methodological limitations of the study. Section two (6.2) discusses the issue of the use of SLA data for the evaluation of typological universals in linguistic typology. Section three (6.3) suggests some theoretical issues arising from the study that warrant future research.

6.1. Methodological Limitations of the Study

Apart from the major findings of the study, there are some perceived sampling limitations of the research that need addressing. The first has to do with the sampling of the subjects on the basis of their English proficiency levels (also used as a variable of the study) at the Monash University English Language Centre. As was mentioned in chapter three (3.1), it was difficult to recruit participants at the lower levels – preliminary and pre-intermediate – because there were fewer international students studying at these levels and even far fewer recruited participants among the targeted population at these levels as compared with the other higher levels. In order to achieve a balanced sample of the study in terms of English proficiency levels, the second round data collection aimed to recruit participants at the higher levels – intermediate, upper-intermediate and advanced. The low-, mid- and high-levels conveniently used in this study were therefore intermediate, upper intermediate and advanced levels respectively rather than the originally attempted range of levels from preliminary to advanced (see 3.1 and note 4 in **Appendix II** for details). As a result, the differences between subjects from different English proficiency levels were not always very big and some results in this respect are even indiscriminate.

Relating to the above sampling factor, different L1 groups, used as another variable of the study, could not be controlled and were therefore unevenly represented in the sample. Seven L1 groups were classified in the study with varying number in each group across different tasks. For example, the Mandarin-speaking group had the largest number of subjects (18) and the Indonesian-speaking group the smallest number (3) across all the

tasks. In view of the unbalanced representation of L1 groups, the results examined via this variable across various tasks should therefore be read with caution.

Finally, a sampling limitation can also be found in the design of the Accessibility Hierarchy-related items in the grammaticality judgment task (see 4.2.3). There were only five task items designed to test the four relativised positions – SU, DO, IO and GEN on the Accessibility Hierarchy with two task items testing DO position (one of them was originally designed as a postnominal RC in contrast to a prenominal RC (S11) in the task). The inclusion of task items for testing OBL and OCOMP positions and also two task items for testing each of the relativised positions would have led to the data in this task being more sufficiently constituted and therefore more comparable to the results of other tasks in this regard. The lack of sufficient data in the examination of the subjects' performances via factors of different English proficiency levels and first languages, may have contributed to some exceptions to the Accessibility Hierarchy in this task, whereas in the repetition and sentence combination tasks in which there were relatively more sufficient data, there were no exceptions under the same examinations.

6.2. Use of SLA Data for the Evaluation of Typological Universals

The use of SLA data for the evaluation of typological universals is significant on two grounds. First, typological universals, which are formulated on the basis of the examination of primary languages of the world, will be more predictive if they are upheld in secondary languages such as interlanguages. If these typological universals are indeed motivated by the processing constraints of the human mind, these constraints should logically apply not only to primary languages but also to secondary languages such as interlanguages since human beings resort to the same processing mechanism in the acquisition of languages, be they primary or secondary. Studies of secondary languages following a typological approach can also inform the study of language universals. The typological universal approach, amongst other approaches, has enriched the inquiries in the field of SLA and the present research is an example of the application of this approach in interlanguage studies.

For this study, the interlanguage data collected from L2 learners of English with a variety of L1-speaking backgrounds forms the basis for testing a number of typological universals. On the one hand, the typological universals are overwhelmingly supported by the interlanguage data of the study, no matter whether these universals have been tested widely such as the Accessibility Hierarchy (Keenan and Comrie 1977), less widely such as those of question formation (Greenberg 1963), or tested for the first time in this study such as those of filler-gap constructions (Hawkins 1999). This is quite significant because most of these universals hold for the interlanguage of the L2 subjects in this study irrespective of their different English proficiency levels and different first languages they speak. On the other hand, despite the interlanguage variability in terms of the differences between individual performances in different types of tasks of the study, common interlanguage syntactic features are found to exist in the way predicted by the typological universals under examination.

Importantly, the subjects' performances have been shown to support the universals in most cases with the exceptions identified. The small number of counterexamples in this study, though not discrediting the overwhelming validity of the universals tested, can also inform these universals nonetheless. When faced with counterexamples, there are three possibilities open (cf. Comrie 1981: 158). The three options for addressing the counterexamples in this study are: 1) the universal as stated is defective in some way; 2) the universal is not relevant to certain structures; or 3) the universal does not apply to secondary languages such as interlanguages. We now examine these cases respectively.

1) The universal as stated is defective in some way – either the universal is a tendency rather than an absolute or counterexamples suggest that it should be modified or abandoned.

I do not believe I have discovered any cases such as this in this study. In Greenberg's (1963) 30-language sample, both implicational universals – “Yes/no Inversion implies *Wh* Inversion” and “*Wh* Inversion implies *Wh* Fronting” are absolute. The one counterexample to the universal “Yes/no Inversion implies *Wh* Inversion” in this study as

well as in Eckman et al (1989), as compared with the overwhelming confirming evidence in both studies, does not suggest that this universal is not absolute, let alone modified or abandoned. The exception in this study has more to do with the idiosyncratic features of interlanguage development of one subject (M7), whose performances were found to violate the universal in the first round collected data but not in the second round.

2) The universal may not be relevant to certain structures, which, on the surface, seem as though they ought to be subject to the particular universal.

I argue here, that the English possessive 's construction which is found in this study to be a counterexample to the Accessibility Hierarchy, may be regarded as a non-canonical genitive construction, and therefore not be subject to the Accessibility Hierarchy.

Apparent exceptions to the Accessibility Hierarchy relating to the subjects' performances on relative clauses which relativise on GEN position were found in the repetition, sentence combination and grammaticality judgment tasks in this study (see 4.2.1, 4.2.2, and 4.2.3). Specifically, there are more instances of (*near*) *correct repetition* and of *predicted RC embedding* on GEN position than on DO, IO/OBL, and OCOMP positions in the repetition and sentence combination tasks respectively, and there are more instances of judging relativisation on GEN position as more acceptable than relativisation on any other position on the Accessibility Hierarchy in the grammaticality judgment task.

These findings are consistent with some similar findings in the literature such as in Gass (1979), and Li and Li (1994), both of which tried to account for their counterexamples from different perspectives. Gass (1979) suggested two possible accounts for her counterexamples in terms of perceivable salience of the English invariant RC marker *whose* and her L2 subjects' specific interpretation of the RC marker and the noun following it as a whole unit, which covers either SU or DO positions (positions high on the Accessibility Hierarchy). Li and Li's explanation for their counterexamples is less convincing, which is similar to Gass's first proposed explanation, but less accurate (see note 24 for details).

Keenan and Comrie (1977) formulated the Accessibility Hierarchy on the basis of a 49-language sample in which about 40 languages have one means or another to relativise on GEN position. Two thirds (27 languages) of the 40 languages demonstrate the preference for the strategy of pronoun retention for relativisation on GEN position (a position that exemplifies the most frequent use of this strategy compared to any other relativised position).⁵⁵ Keenan and Comrie (1977: 92) suggested that the pronoun retaining strategy is used "in proportion to the difficulty of the position being relativized"; that is, the lower the relativised position down the hierarchy, the more the languages using the strategy will be found as long as relativisation is allowed. English is among the languages that do not use the pronoun retention strategy but allow relativisation on GEN position. For example, although the English genitive RC marker *whose* is functionally similar to the French *dont* and the Swedish *vars*, the latter two are, nevertheless, preferably replaced in practice by an alternative relativisation on SU position in a causative construction (Keenan and Comrie 1977: 91). It thus seemed that the freedom with which English allows relativisation on GEN without a pronoun copy via the RC marker *whose* is different from that in the majority of the world languages.

It is crucial here to recall that English has two ways of forming possessives, one with the possessive *'s*, and one with the *of*-construction. I would argue that it is the latter rather than the former that may be more comparable to typical genitive constructions in the world's languages.

Croft (1990: 28-32) enumerated ten typologically different genitive constructions, among which *Linker* is a "rare strategy".⁵⁶ He argues that "The English genitive *-s* is a unique possessive morpheme", and "would probably best be analyzed as a linker, because it

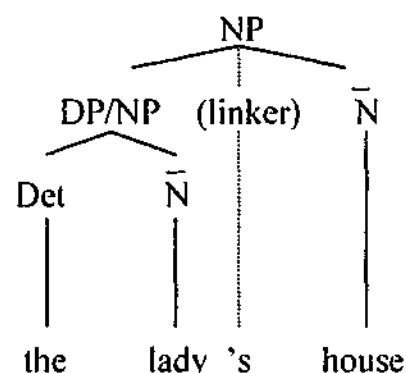
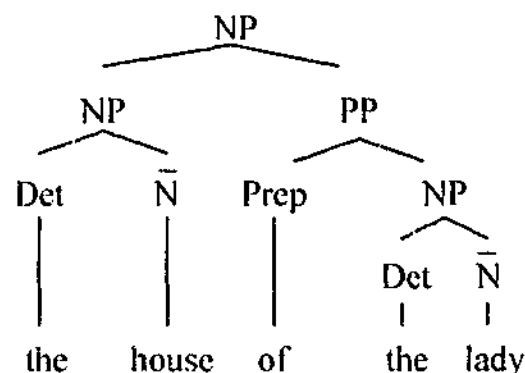
⁵⁵ Refer to Tables 1 and 2 in Keenan and Comrie (1977: 76-9, 93) for details. Keenan and Comrie (1977: 93) listed 26 languages that have pronoun retention in relative clauses, but Hawkins (1999: 258) mentioned 27 languages, adding Tongan, an ergative language, in the list.

⁵⁶ "*Linker*: A rare strategy, found chiefly in adjectival, genitive and other noun-noun constructions, is to use an additional morpheme called a 'linker.' The linker is normally invariant marker used for modifiers in noun phrases but not in predicate-argument (verb-noun-phrase) relations in clauses. The linker is invariant or contrasts only with simple juxtaposition, and functions merely to link the possessor and head noun

contrasts primarily with noun-noun juxtaposition, it does not occur in the same syntactic position as the English prepositions, and is not used for verb (predicate)-argument NP relations" (Croft 1990: 32). In genitive constructions, it is the *case marker* and the *adposition* – "the independent counterpart of the case marker" (Croft 1990: 30) – rather than the *linker* that are more typologically attested.

The structures of the possessive morpheme 's in English (as in *the lady's house*) and of-genitive construction (as in *the house of the lady*) might be illustrated in the tree representations in (1).⁵⁷

(1) a. possessive 's

**b. *of*-genitive**

The structure in 1b involves two full NPs and a preposition, whereas the structure in 1a has only one full NP (It is not possible to have independent determiners for both nouns in the English possessive 's construction). When relativisation on GEN position occurs in English, there are also two alternatives involving *whose* or *of-relative pronoun* as shown in (2).⁵⁸

- (2) a. The lady *whose* house we bought teaches Spanish.
b. The table, the top *of which* we scratched, lost a lot of its value.

grammatically. The linker may be bound to one (or both) of the constituents, or function as an independent particle." (Croft 1990: 32)

⁵⁷ I believe that within the Chomskyan approach, 's' is not assigned to a category but treated in some places as a kind of spelling out feature.

⁵⁸ The invariant English RC marker *whose* can be similarly analysed as “who + a linker -s” or “who’s”, which is exemplified in the subjects’ performance in the sentence combination task.

Similarly, it may well be that *of which* in 2b rather than *whose* in 2a is more comparable to the typical structure of GEN relativisation in the world's languages, and more comparable to the relativised GEN position on the Accessibility Hierarchy as well. Because GEN position, low on the Accessibility Hierarchy, involves processing complexity in gap identification, the majority of the languages that allow relativisation on GEN position at all, resort to the pronoun retaining strategy by avoiding the gapping strategy (Hawkins 1999). English does not have the pronoun retaining strategy for relativisation on GEN; instead, it has the gapping strategy by using uniquely an invariant RC marker *whose*, which is more frequently used than the '*of-relative pronoun*' structure and is probably perceivably salient (Gass 1979).⁵⁹ In Hawkins' terms, the filler-gap domain with a gap (**whose**) in the possessive 's construction such as 1a, contains fewer nodes to be processed than the filler-gap domain with a gap (**of-relative pronoun**) in the *of*-genitive construction such as 1b, and could therefore be expected to be more accessible. If the examination of relativisation on GEN position was restricted to the '*of-relative pronoun*' structure only, the results for English interlanguage would predictably support the Accessibility Hierarchy without exceptions.

Provided the above argumentation holds, we can tentatively conclude that since the English possessive 's construction is typologically unusual, the L2 learners' performance on relativisation on GEN position in English (as shown by the results of this study as well as those in Gass 1979, and Li and Li 1994) is irrelevant to the examination of the Accessibility Hierarchy. In that case, the exceptions relating to English relativisation on GEN position are not damaging to the validity of the Accessibility Hierarchy nor to its underlying processing motivated assumption.

3) A third way to approach apparent counterexamples to proposed universals is to consider whether: *certain universals may not apply to secondary languages such as interlanguage at certain stages of acquisition due to the semantic or syntactic complexity they involve.*

⁵⁹ The use of *whose* and *of-relative pronoun* also has to do with animacy. Compared with (2), "The lady, the house of whom we bought teaches Spanish" seems awkward in English, while "The table whose top we scratched lost a lot of its value" seems acceptable.

The exceptions to the Head Noun Phrase Hierarchy may be explained in terms of the semantic complexity involved. It is proposed that the exceptions to the Head Noun Phrase Hierarchy (Hawkins 1999) are due more to the complexity of the use of English articles and the level of competence of L2 learners than to the implicational universal itself.

The English articles are very complex in that the appropriate or correct use of them involves comprehending and discriminating between a whole set of semantic features including definiteness, specificity and anaphoric reference as well as syntactic contexts such as number. For L2 learners, especially for those whose first languages do not have a category of articles, it seems difficult for them to use the English articles appropriately especially in contexts that require fine discrimination. For example, in the English sentence "The book is on the table", the definiteness of *book* and *table* is expressed with the preceded definite article *the*, but in Chinese the definiteness is expressed with bare nouns *shu* (the book) and *zhuozi* (the table). Therefore, it is likely that the subjects in this study have not fully acquired the set of semantic and syntactic features of the English articles, let alone in a context involving more complex structures of filler-gap dependencies, which contributes primarily to the counterexample to the Head Noun Phrase Hierarchy.

The case of indiscriminate responses of the subjects to both the Clause Embedding Hierarchy and the Valency Completeness Hypothesis (see 4.3.7) appears to be a function of the structural complexity of *wh*-islands and sentential subjects in English that are beyond the competence of the subjects in some way. It is not the case that ESL learners have the same competence as L1 speakers. What has been observed in this study is that the subjects have differential mastery of structures that accord with the relative complexity of structures as characterised by the relevant universals. It is perhaps not surprising that there is a limit to the degree of complexity second language learners can process.

6.3. Some Theoretical Issues Arising from the Study That Warrant Future Research

On the basis of the findings of this study, some possible areas for future research are suggested. First, the implicational universals pertaining to question formation (Greenberg 1963) were predominantly supported by this study as well as the study of Eckman et al (1989). However, there was one exception to the universal "*Yes/no Inversion* implies *Wh Inversion*" in both studies, and there was a ceiling effect in the use of initial *wh* words in the universal "*Wh Inversion* implies *Wh Fronting*" among all the subjects in both studies as well. The subjects in both studies were L2 learners of English at the English proficiency levels well above beginner or preliminary levels. It might therefore be interesting to see whether more exceptions would arise and whether the ceiling effect of the use of initial *wh* words in *wh*-questions will persist in future studies using beginning L2 learners of English as subjects.

The Accessibility Hierarchy (Keenan and Comrie 1977) was supported by all types of data in this study. Interestingly, although both the Accessibility Hierarchy and the Absolutive Hypothesis (Fox 1987) hold in the naturalistic speech and writing, stronger support for the former is found in the written essays and stronger support for the latter comes from the conversation (see 4.2.4). To gain a better understanding of the extent to which each of the two implicational universals holds for what type of interlanguage data, it is suggested that future research include a variety of tasks apart from naturalistic data to examine the relativisation on the arguments – S, A and P. The discussion pertaining to the exceptions to the Accessibility Hierarchy (see 6.2) suggests that it will be theoretically interesting to see whether exceptions will still arise in studies that use the '*of-relative pronoun*' structure instead of the *whose* structure for relativisation on GEN position.

Finally, since this study is the first tentative attempt (as far as the writer knows) to test the recently motivated implicational hierarchies/hypotheses of Hawkins (1999) in L2 settings, more research is warranted. Furthermore, it will be interesting to see whether the processing account in term of filler-gap domains associated with various types of

structural complexity, which explains L2 phenomena adequately in this study, is further supported in second language acquisition studies to come.

APPENDICES

Appendix I Table of Universals Tested, Data Collected and Methods Used in the Study

Implicational Universals*	Data	Methods
Universal 11 of Greenberg (1963): If yes-no inversion, then Wh inversion; if Wh inversion, then initial Wh words/phrases	Naturalistic oral production	Conversation task (including an interview & a role-play)
Accessibility Hierarchy SU > DO > IO > OBL > GEN > OCOMP (Keenan and Comrie 1977)	Oral & written; comprehension & production	Tasks of conversation, repetition, sentence combination, & grammaticality judgment, and collected written essays
Filler-gap dependencies (Hawkins 1999) a. Clause Embedding Hierarchy	Comprehension	Grammaticality judgment task
b. Subordinate Gap/No Gap Hierarchy	Comprehension	Grammaticality judgment task
c. Hierarchy for Complementisers (Kluender 1992)	Comprehension	Grammaticality judgment task
d. Bridging Verb Hierarchy	Comprehension	Grammaticality judgment task
e. Head Noun Phrase Hierarchy	Comprehension	Grammaticality judgment task
f. Valency Completeness Hypothesis	Comprehension	Grammaticality judgment task

* For detailed illustration of these universals, refer to chapter 2.

Appendix II

Profiles of the Subjects

No.	Code Name*	Gender	Age	Country of Origin	First Language	Years of Learning English	Months of Stay in Australia	Class Level****
1	M1	F	25	Taiwan	Mandarin	13***	2	b-GE4C
2	M2a/ M2b**	F	37	China	Mandarin	24	1	c-GE5/ ETP6A
3	M3	M	23	Hong Kong	Mandarin	20	1 & 1/2	c-GE5
4	M4	M	29	China	Mandarin	15	3	b-GE4B
5	M5a/M5b	M	26	China	Mandarin	13	3	b-GE4B/ETP6A
6	M6	M	20	China	Mandarin	7	2	b-GE4C
7	M7a/M7b	F	25	China	Mandarin	12	3	d-ETP6/F. offer
8	M8	M	25	Taiwan	Mandarin	12	12	b-GE4B
9	M9	M	19	China	Mandarin	7	2	d-EPT6A
10	M10	F	26	China	Mandarin	14	1	c-ETP5A
11	M11	F	23	China	Mandarin	11	4	d-EPT6A
12	M12	F	23	China	Mandarin	11	2	c-ETP5A
13	M13	F	20	China	Mandarin	6	3	c-ETP5A
14	M14	M	24	China	Mandarin	12	1	c-EPT5A
15	M15	F	23	China	Mandarin	13	1	c-ETP5A
16	J1	F	28	Japan	Japanese	15	1 & 1/2	b-GE4C
17	J2	F	20	Japan	Japanese	8	1	a-GE3B
18	J3	M	20	Japan	Japanese	8	1 & 1/2	b-GE4C
19	J4	M	20	Japan	Japanese	8	1	b-GE4B
20	J5	M	20	Japan	Japanese	10	1 & 1/2	b-GE4B
21	J6	F	20	Japan	Japanese	9	2	b-GE4B
22	J7	F	20	Japan	Japanese	8	1	a-GE2
23	J8	F	21***	Japan	Japanese	9***	1	c-GE5
24	J9	M	21***	Japan	Japanese	9***	2	b-GE4B
25	J10	M	27	Japan	Japanese	14	1	b-GE4B
26	J11	F	20***	Japan	Japanese	8***	4	b-GE4C
27	J12	F	22	Japan	Japanese	9	2	c-GE5A
28	K1	F	26	Korea	Korean	13	5	c-GE5
29	K2	M	20***	Korea	Korean	8***	3	b-GE4B
30	K3	M	23	Korea	Korean	10	2	b-GE4B
31	K4	M	27	Korea	Korean	14	6	c-GE5A
32	K5	M	21	Korea	Korean	8	7	c-GE5A
33	K6	M	23	Korea	Korean	10	3	b-GE4B
34	K7	M	20	Korea	Korean	7	2	d-GE6A
35	K8	M	21	Korea	Korean	8	3	c-EPT5A
36	K9	M	23	Korea	Korean	10	4	c-GE5A
37	V1a/V1b	M	25	Vietnam	Vietnamese	4	4	a-GE3B/ETP6A
38	V2	F	19	Vietnam	Vietnamese	13	4	c-GE5
39	V3	F	23	Vietnam	Vietnamese	13	4	b-GE4B
40	V4	M	23	Vietnam	Vietnamese	2-3	4	d-EPT6A
41	V5	F	41	Vietnam	Vietnamese	2-3	4	d-EPT6A
42	V6	F	27	Vietnam	Vietnamese	14	1	d-ETP6A
43	V7	M	25	Vietnam	Vietnamese	1	6	d-EPT6A
44	T1	F	23	Thailand	Thai	10	4	d-GE6A
45	T2	F	23	Thailand	Thai	13	3	c-EPT5A
46	T3	F	23	Thailand	Thai	10	7	d-GE6A
47	T4	F	21	Thailand	Thai	8	8	d-GE6A
48	T5	M	23	Thailand	Thai	10	7	d-GE6A
49	T6	F	23	Thailand	Thai	10	5	c-ETP5A
50	T7	F	23	Thailand	Thai	11	5	c-ETP5A
51	Int	M	19	Indonesia	Indonesian	10	2	c-GE5

52	In2	M	23	Indonesia	Indonesian	10	3	c-EPT5A
53	In3	F	20	Indonesia	Indonesian	9	3	d-ETP6A
54	S1a/S1b	M	33	Colombia	Spanish	6	2	a-GE2/ GE4B
55	S2	F	37	Colombia	Spanish	25	4	c-GE5A
56	F1	F	19	Switzerland	French	6	6	d-ETP6A
57	It1	F	26	Switzerland	Italian	5	5	d-EPT6A
58	B1	F	29	Bangladesh	Bangla	19	4	d-EPT6A

NOTE:

- *: The code name for each subject is composed of a letter (or two) indicating his/her first language, and a number indicating the order he/she is assigned within his/her first language group. Those subjects who participated in none of the tasks but the written essay collection were not listed here (see **Appendices VIII** and **XI** for their performances in the written essays).
- **: Subjects who participated in both rounds of data collection with *a* referring to the first round and *b* to the second round.
- ***: Approximate age entered when raw data was deficient or years of learning English approximated when numerical figures were not specifically stated (e.g., since Junior High School).
- ****: There were seven class levels for the subjects: GE 2 (N = 2), GE 3 (N = 2), GE 4 (N = 18), GE 5 (N = 11), GE 6 (N = 5), ETP 5 (N = 10) and ETP 6 (N = 15) (Those who participated in both rounds, were counted separately).
- According to their corresponding levels, the seven class levels were re-grouped into four levels:
- preliminary to pre-intermediate: GE 2 & 3 (N = 4)
 - intermediate: GE 4 (N = 18)
 - upper-intermediate: GE 5 & ETP 5 (N = 21)
 - advanced: GE 6 & ETP 6 (N = 20)
- For the purpose of the analysis and reporting of the results in the main text, these were further reduced to three levels by collapsing levels *a* and *b* into one, conveniently labeled as low level (*a* and *b*), mid level (*c*), and high level (*d*) respectively.

Appendix III Tasks for Data Collection (Approx. 35 minutes)

I. Conversation

A. Interview (Approx. 5 – 8 minutes)

The native English speaker will talk with each participant following the guidelines set below:

- a. Exchange greetings and settle the participants down.
- b. Ask basic questions about personal background information (age, country of origin, native language, years of studying English, years of staying in Australia, etc.)
- c. Ask the participants to describe their studies (importance of learning English, learning methods, difficulties and problems, differences between English and their native languages, etc.), interests/hobbies, campus life, work and/or some other topic that they can talk about in detail.
- d. Ask the participants about their future plans after their studies at the Language Centre.
- e. Conclude the interview and go to the next task.

B. Cued Elicitation of Questions (Approx. 3 – 5 minutes)

Information about *wh* and *yes/no* questions will be elicited in the following 'information gap' task. Each participant will be given one of the two cued cards to start a conversation with the native speaker and gain information as specified by the cues from her.

Card One: Buying a secondhand TV set: You want to buy a secondhand TV set. The native speaker is in charge of a shop that sells secondhand TV sets. Ask her for information about the TV set you want by using *yes/no* questions and/or questions with words like *who*, *what*, *where*, *why*, *how*, *how much*, and so on.

Find out about: brand
 size
 price
 remote control
 delivery
 warrantee

Card Two: Melbourne: You are interested in going sightseeing in Melbourne. The native speaker comes from Melbourne. Ask her for information about Melbourne by using *yes/no* questions and/or questions with words like *who*, *what*, *where*, *why*, *how*, *how much*, and so on.

Find out about: weather
 size
 places of interest
 shops
 transport
 entertainment

II. Repetition (Approx. 5 minutes)

Ask participants to repeat the following sentences after the native speaker recorded on the tape.

1. The dog which Penny bought today is very gentle
2. The man who Neil is as rich as gave less.
3. The girl who got the answer right is clever.
4. The boy whose brother tells lies is always honest.
5. The girl who Sue wrote the story with is proud.
6. The man who Ann gave the present to was old.
7. The girl whose friend bought the cake was waiting.
8. The food which Chris paid the bill for was cheap.
9. The dog which Mary taught the trick to was clever.
10. The letter which Dick wrote yesterday was very long.
11. The boy who Mike writes better than was listening.
12. The boy who told the story was very young.

III. Sentence Combination (Approx. 5 minutes)

Read the following pairs of sentences and combine each pair into one sentence containing a relative clause like the following example.

Example: The boy ran to school.

The boy was singing. → The boy who was singing ran to school.

1. The policeman caught the thief. The thief stole the car.
2. The bookstore changed its name. The bookstore sells science books.
3. The dog ate the bone. The boy found the bone.
4. The speech impressed the audience. The chairman made the speech.
5. The girl saw the actor. The girl sent a letter to the actor.
6. The boy broke his words. The girl lent her car to the boy.
7. The detective questioned the man. You work with the man.
8. The dog won a prize. Sue told her mother about the dog.
9. The manager sacked the man. The man's negligence caused the accident.
10. My son saw the lady. We bought her house last week.
11. The boy stole the jewels. His father teaches law.
12. The patient committed suicide. The doctor regarded his disease as incurable.
13. The man won the race. You run as fast as the man.
14. The man saw the lady. His wife is less beautiful than the lady.

IV. Grammaticality Judgment (Approx. 5 – 10 minutes)

Read the following sentences and circle one of the choices at the end of each sentence.

V : acceptable English

X : unacceptable English

? : not sure

- | | | | |
|---|---|---|---|
| (1) The dog that is barking at the child is a poodle. | V | X | ? |
| (2) The lady whose purse was stolen went to the police station. | V | X | ? |

(3)	The problem that we talked about it for a long time has been solved now.	V	X	?
(4)	Tony lost the dictionary who you borrowed last week.	V	X	?
(5)	The boy who Mary sent a postcard to is her brother.	V	X	?
(6)	The man hit the boy who you saw in the supermarket.	V	X	?
(7)	The lady talked to John yesterday is his mother.	V	X	?
(8)	The student was very happy whose paper got the highest score.	V	X	?
(9)	What did she tell John about?	V	X	?
(10)	What did John doubt whether she would win?	V	X	?
(11)	He interviewed the man yesterday has come.	V	X	?
(12)	Has Mary told you to do what?	V	X	?
(13)	What did John doubt if she would win?	V	X	?
(14)	Who did you see a picture of?	V	X	?
(15)	Do you think who the mayor will meet?	V	X	?
(16)	What did John doubt that she would win?	V	X	?
(17)	Who did you see the picture of?	V	X	?
(18)	How angry did Mary say that John was?	V	X	?
(19)	The book she bought last week is missing.	V	X	?
(20)	What did John doubt the truth of?	V	X	?
(21)	How angry did Mary say softly that John was?	V	X	?
(22)	What did you wonder how they would bake?	V	X	?
(23)	Who did you see John's picture of?	V	X	?
(24)	Who do you know the professor that taught?	V	X	?
(25)	The person that you hoped to see is Harry.	V	X	?
(26)	How angry did Mary whisper that John was?	V	X	?
(27)	Who did it surprise Sue that Mary disliked?	V	X	?
(28)	What did you wonder how to bake?	V	X	?
(29)	The person that you hoped that you would see is Harry.	V	X	?
(30)	What did the title of amuse John?	V	X	?
(31)	What do you regret the fact that he stole?	V	X	?
(32)	What did to read fascinate Sue?	V	X	?
(33)	The student who you know the professor that taught is Harry.			
(33')	The person that you thought that Bill said Harry wanted to see has gone overseas.*	V	X	?
(34)	Who did that Mary disliked surprise Sue?	V	X	?
(35)	What did John read the title of?	V	X	?
(36)	What did it fascinate Sue to read?	V	X	?

* Task item (33) was used in the first round of data collection and task item (33') in the second round.

Appendix IV Counts of Instances in Features of Question Word Order

No.	Code Name of Subjects	Total		Feature A (Wh Fronting)	Feature B (Wh Inversion)	Feature C (Yes/no Inversion)
		WHQ	YNQ			
1	M1	3	6	3	3	5
2	M2b	4	12	4	2	6
3	M3	2	6	2	2	4
4	M4	3	5	3	3	4
5	M5b	3	4	3	3	4
6	M6	1	6	1	1	4
7	M7a	3	6	3	1	5
8	M7b	1	9	1	1	8
9	M8	3	2	3	3	2
10	M9	5	4	5	5	1
11	M10	3	7	3	3	5
12	M11	1	7	1	1	7
13	M12	2	3	2	2	3
14	M13	2	17	2	2	17
15	M14	4	3	4	4	3
16	M15	3	7	3	3	3
17	J2	2	6	2	2	5
18	J3	5	6	5	4	6
19	J4	4	7	4	4	5
20	J6	2	2	2	2	2
21	J8	4	14	4	4	13
22	J10	1	6	1	1	3
23	J11	1	3	1	1	2
24	K1	9	9	9	8	5
25	K2	2	2	2	2	1
26	K3	3	7	3	3	6
27	K4	4	4	4	4	4
28	K5	2	5	2	2	4
29	K6	3	7	3	3	6
30	K7	3	6	3	3	5
31	K8	2	9	2	2	7
32	V1a	4	4	4	4	4
33	V4	6	1	6	5	1
34	V5	3	8	3	3	8
35	V6	1	6	1	1	5
36	V7	2	6	2	2	5
37	T1	3	8	3	3	6
38	T2	3	9	3	3	5
39	T3	4	4	4	3	4
40	T4	2	3	2	2	2
41	T5	3	7	3	2	4
42	T6	4	5	4	4	4
43	T7	2	10	2	2	5
44	In1	2	7	2	1	4
45	In2	4	8	4	4	6
46	In3	6	6	6	6	4
47	S1a	3	8	3	2	4
48	S1b	4	6	4	4	5
49	S2	3	2	3	2	1
50	F1	4	7	4	4	7
51	It1	4	1	4	4	1
52	B1	9	1	9	8	1
Total				166	153	240

Note: In the table, Column 1 shows the sequential number of subjects (50 in all with M7 and S1 in both rounds). Column 2 lists code names of the subjects with letters indicating the first language (B = Bangla, F = French, I = Indonesian, It = Italian, J = Japanese, K = Korean, M = Chinese, S = Spanish, T = Thai, V = Vietnamese) and digital figures indicating the sequential number of the subjects within that particular first-language group. Column 3 provides the total number of *wh* and *yes/no* questions obtained from each subject. Columns 4-6 provide figures involving the three features respectively. Column 4 shows the number of instances of Initial *wh* words/phrases, Column 5 shows the number of instances of *wh* inversion, and Column 6 shows the number of instances of *yes/no* inversion for each subject.

Appendix V Definitions of Features for the Repetition Task (A)

1. **Correct repetition:** Sentences are considered as correct if subjects miss or misuse no more than two words which do not occupy the positions of S, V, O or Complement. In addition, words with incorrect number, tense, aspect, or agreement, and omitted non-obligatory RC markers are disregarded.
e.g. The dog which Penny bought today is gentle.
The dogs Penny bought today were very gentle.
2. **Near correct repetition:** Sentences are considered as nearly correct if subjects misuse a verb, miss or misuse one word which occupies the positions of S, O or Complement (no more than two missed or misused words in total) or if the verb is in the passive instead of the active voice. In addition, words with incorrect number, tense, aspect, or agreement, and omitted non-obligatory RC markers are disregarded.
e.g. The dog which/that Penny put/bring today is very gentle.
The dog which Penny bought today is very...
3. **Failed repetition:** Classified in this feature are examples that include non- or partial repetition, and attempted repetition that cannot be classified according to the other categories.
e.g. The dog which Penny today is very gentle.
The dog with Penny today is very gentle.
Dog is Penny today is very gentle.
4. **Incorrect RC marker:** Examples classified in this feature includes sentences in which subjects have used an incorrect RC marker such as 'who' instead of 'whose' or 'who...with'.
e.g. The girl who friend bought the cake is waiting.
The girl who Sue wrote the story is proud.
5. **Different meaning:** Sentences which have the same structure of relativisation as the original sample sentences, have quite different meaning due to choice of different lexical items.
e.g. A bog I bought yesterday is very dangerous.
The bog which Penny paid today is gentle
6. **Omission of obligatory RC marker:** Classified in this category is a sentence which omits an obligatory relative clause marker, the insertion of which would make the sentence either as an instance of *correct repetition* or one of *near correct repetition*.
e.g. The girl got the answers right is clever.
The dog is taught a trick to is clever.

- 7. Different relativisation:** Sentences which relativise on different positions from those of the sample sentences. Such examples may include sentences that are quite similar in meaning to the original, and sentences that have a significantly different meaning.

e.g. The dog which born Jenny today is very gentle.
The dog which played today is very gentle.

- 8. Topic-comment structure:** The sample sentence is expressed in a topic-comment structure in which the topic is relativised on.

e.g. The girl who wrote this story, she is proud.
The boy who Mike write better than, he is listening.

- 9. Resumptive words:** Classified in this category is a relative clause in which a resumptive pronoun or noun is used.

e.g. The girl who his friend bought the cate is waiting.
The letter which was Dick wrote it was long.

Individual Performances in the Repetition Task (B)

No	Code Name	SU		DO		IO		OBL		GEN		OCOMP	
		S3	S12	S1	S10	S6	S9	S5	S8	S4	S7	S2	S11
1	M1	3	3	6	3	3	3	3	3	3	3	3	7
2	M2a*	1	1	3	3	7	3	3	3	2	1	7	3
3	M2b	1	1	3	2	2	7	2	3	1	1	7	3
4	M3	6	1	7	3	7	3	3	3	4	2	3	3
5	M4	3	2	3	1	3	7	3	3	1	7	3	3
6	M5a	6	1	3	3	3	7	3	3	3	3	3	7
7	M5b	6	1	3	3	7	3	3	4, 7	3	3	7	1
8	M6	1	1	1	7	7	3	3	3	1	7	3	3
9	M7a	1	1	7	1	2	7	7	7	1	1	7	7
10	M7b	1	1	1	1	2	1	7	1	1	1	7	3
11	M8	1	1	3	3	7	7	4	7	7	7	7	7
12	M9	1	1	1	3	5	3	1	3	3	7	3	1
13	M10	1	1	1	3	3	3	3	7	2	3	3	3
14	M11	1	1	3	7	7	3	7	3	7	7	7	3
15	M12	2	1	7	3	3	3	3	3	7	3	3	3
16	M13	1	1	3	3	5	7	1	1	7	4	7	7
17	M14	1	1	3	1	2	3	7	3	1	4, 9	7	3
18	M15	1	1	1	1	7	3	7	3	1	1	7	7
19	J1	6	2	3	3	3	3	3	3	3	3	3	3
20	J2	1	1	1	3	4	3	3	3	1	7	3	2
21	J3	3	1	3	3	3	3	7	3	3	7	3	7
22	J4	1	1	1	3	4	7	4	3	1	2	3	3
23	J5	2	1	3	3	3	3	3	3	7	7	7	3
24	J8	2	1	3	3	3	3	3	3	1	3	7	3
25	J9	1	1	1	1	3	3	3	7	7	7	7	3
26	J10	1	1	3	3	3	3	3	3	3	3	3	3
27	J11	3	1	3	2, 9	3	3	3	3	7	7	7	3
28	J12	6	1	3	3	3	3	3	3	3	3	3	3
29	K1	1	2	3	3	7	3	3	3	7	3	7	3
30	K2	1	1	3	3	3	6	7	3	1	2	3	3

31	K3	3	1	3	3	7	3	3	3	3	7	3	3
32	K4	3	1	3	3	3	3	3	3	3	3	3	3
33	K5	3	2	3	3	7	3	3	3	3	3	3	3
34	K6	3	1	3	3	7	3	3	3	7	3	3	3
35	K7	1	1	3	1	4	1	4,8	2	1	2	7	3
36	K8	8	8	1	3	7	7,8	7,8	1	4, 8, 9	3	3	7
37	K9	2	3	1	3	7	7	7	7	3	3	3	3
38	V1a	6	1	5	2	3	3	3	3	3	7	3	3
39	V1b	2	1	3	3	3	3	7	3	3	1	7	3
40	V3	1	2	2	3	4	3	3	3	7	1	7	7
41	V4	1	1	1	1	3	3	3	7	7	2	3	3
42	V5	1	1	5	3	1	3	3	3	3	1	3	8
43	V6	3	8	3	5,8	3	3	3	3	3	3	3	3
44	V7	1	1	1	3	3	3	3	7	3	3	3	3
45	T1	1	2	3	3	3	3	3	7	4	7	7	2
46	T2	1	1	3	3	7	3	3	2	2	3	3	3
47	T3	1	1	1	3	3	5	3	3	7	2	7	7
48	T4	1	1	2	1	2	1	7	3	4	1	3	3
49	T5	1	1	1	3	3	3	3	7	1	4	3	3
50	T6	1	3	3	3	3	3	3	3	7,8	3	3	3
51	T7	1	1	3	1	7	3	3	3	1	3	3	3
52	In1	1	1	3	3	1	5	4	3	7	7	3	3
53	In2	6	1	3	2	3	3	3	3	3	3	3	3
54	In3	1	1	2	5	1	7	7	3	1	1	7	3
55	S1a	3	3	3	3	3	3	3	3	3	3	3	3
56	S1b	6	1	3	3	3	3	3	3	3	3	3	3
57	S2	3	1	3	3	3	3	3	3	3	3	3	3
58	F1	1	1	2	4	1	4	1	1	1	1	3	4
59	It1	1	1	1	4	7	4,7	4	3	3	7	3	3
60	B1	1	1	2	1	1	3	3	3	1	1	3	3

* Subjects who participated in both rounds of data collection with *a* referring to the first round and *b* to the second round.

Appendix VI

Definitions of Features for the Sentence Combination Task (A)

1. **Predicted RC embedding:** Classified in this feature are sentences that are combined in accordance with the requirement – the combined sentence should contain a relative clause with the first sentence in a pair becoming a matrix clause and the second embedded in the first as a relative clause. Words with incorrect number, tense, aspect, or agreement, and omitted non-obligatory RC markers are disregarded.
e.g. The dog ate the bone the boy found.*
2. **a. Relativisation on a similar position:** Sentences classified in this feature include those that relativise on a position different from the predicted relativised position but ranked as the **same** on the hierarchy. Words with incorrect number, tense, aspect, or agreement, and omitted non-obligatory RC markers are disregarded.
e.g. The boy found the bone that the dog ate.

b. Relativisation on a higher position: Sentences classified in this feature include those that relativise on a (different) position that is **higher** than the predicted relativised position on the hierarchy. Words with incorrect number, tense, aspect, or agreement, and omitted non-obligatory RC markers are disregarded.
e.g. The dog ate the bone which was found by the boy.
The boy found the bone which was eaten by the dog.

c. Relativisation on a lower position: Sentences classified in this feature include those that relativise on a (different) position that is **lower** than the predicted relativised position on the hierarchy. Words with incorrect number, tense, aspect, or agreement, and omitted non-obligatory RC markers are disregarded. No exemplifying examples are found in this category.
3. **No adjacency to the head noun:** Classified in this feature is a relative clause that is separated from its head noun.
e.g. The man saw the lady whose wife is less beautiful than the lady.
The bookstore changed its name that sells science books.
4. **Inaccessible head noun:** Examples classified in this feature include relative clauses whose head nouns are either in possessive case or non-existent.
e.g. The man's negligence caused the accident who was sacked the man.
5. **Incorrect relative markers:** Examples classified in this feature include relative clauses that use incorrect case marking or animacy on a relative pronoun, or omit an obligatory preposition used with a relative marker.
e.g. The bone who found by the boy was ate by the dog.
The manager sacked the man who's negligence caused the accident.

6. **Topic-comment structure:** Sentences classified in this feature contain a topic-comment structure in which the topic is relativised on.
e.g. The boy who found the bone the dog ate it.
The man who saw the lady his wife is less beautiful than the lady.
7. **Resumptive words:** Classified in this feature is a relative clause in which a resumptive pronoun or noun is used.
e.g. The man saw the lady who his wife is less beautiful than her.
The man whose his wife is less beautiful than the lady saw the lady.
8. **Omission of obligatory RC marker:** Classified in this feature is a sentence which omits an obligatory relative clause marker, the insertion of which would make the sentence either as an instance of *predicted sentence combination* or one of *relativisation on a higher position*.
e.g. The bookstore sell science book has changed its name.
Sue told her mother about the dog won a prize.
9. **Other RC-related non-standard forms:** Classified in this feature are errors concerning relative clauses such as double relativisation without a matrix clause, which can not be classified according to category 2, 3, 4, 5, 6, or 7.
e.g. The man whose wife is less beautiful than the lady who he saw.
10. **Failure to supply a relative clause:** Examples classified in this feature include no try, incomplete sentences, unintelligible sentences as well as different types of sentences other than relative clauses.
e.g. The manager sacked the man's accident.
The dog ate the bone after a boy found the bone.

* Examples are cited from the subjects' performances in task items S2, S5, S9, S11 and S14, which are presented as follows along with their respective predicted RC embedding:

- S2: The dog ate the bone. The boy found the bone. → The dog ate the bone that/which the boy found.
- S5: The manager sacked the man. The man's negligence caused the accident. → The manager sacked the man whose negligence caused the accident.
- S9: The man saw the lady. His wife is less beautiful than the lady. → The man saw the lady who his wife is less beautiful than.
- S11: the dog won a prize. Sue told her mother about the dog. → The dog that/which Sue told her mother about won a prize.
- S14: The bookstore changed its name. The bookstore sells science books. → The bookstore that/which sells science books changed its name.

Individual Performances in the Sentence Combination Task (B)

No	Code Name	SU		DO		IO		OBL		GEN-SU		GEN-DO		OCOMP	
		S1	S14	S2	S13	S3	S12	S4	S11	S5	S6	S8	S10	S7	S9
1	M1	1	3	2b	2b, 5	2b, 3	2b	2b	2b	2b, 4	3	2b, 4	2b	2b	2b
2	M2a*	1	1	1	1	2b	2b	1	2b	1	1	1	1	2b	2b
3	M2b	1	1	2a	1	1	1	2b	1	1	1	1**	1	2b	2b
4	M3	1	1	2b	2b	10	2b	2b	2b	10	10	10	2b	10	2b
5	M4	1	1	2a	7	1	7	1	7	1	1	2b	2b	1	7
6	M5a	1	1	2b	1	2b	2b	2b	2b	1	1	2b	2b	7	2b
7	M5b	1	5	1	2b	2b	2b	1	2b	1	10	2b, 4	1**	1	2b
8	M6	1	5	2b	2b	10	2b	2b	2b	1	1	1**	1**	2b	2b
9	M7a	1	3	2a	2b	2b, 5	10	1	2b	1	3	2b, 4	10	3	2b, 3
10	M7b	1	3	2b	2b	1	3	1	2b	1	1	2b, 4	2b	7	1
11	M8	1	2a	2b	2b	2b	2b	1	2b	1	1	2b	2b	1	2b
12	M9	1	1	1	2b	1	2b	1	2b	1	1	1	1	2b	2b
13	M10	1	8	1	2b	10	2b	1	8	10	1	10	10	8	10
14	M11	1	1	1	2b	2b, 5	2b	1	2b	1	1	2b	2b	1	2b
15	M12	1	1	2b	1	2b	2b	1	2b	1	1	1**	1	1	2b
16	M13	1	1	1	1	2b	5	1	1	1	1	1	1	1	1
17	M14	10	10	10	2b	10	10	1	2b	10	1	10	1**	2b	2b
18	M15	1	1	2b	2b	1	1	1	1	1	1	1**	1	2b	2b
19	J1	1	3	10	8	10	8	2b	8	1sp	10	10	10	10	10
20	J2	1	1	1	1	5	1	1	5	1	1	1	1	2b	2b
21	J3	1	10	1	2b	10	10	1	8	1	10	10	10	10	2b
22	J4	1	1	1	1	1	5	1	1	1	1	10	10	1	2b
23	J5	1	1	2b	2b	1	2b	10	8	1	1	2b	1**	2b	2b
24	J8	1	1	1	1	1	1	1	1	1	1	2b	1	2b	2b
25	J9	1	1	2b	2b	2b	2b	2b	2b	10	10	10	2b	2b	2b
26	J10	1	1	1	1	1	1	1	1	1	1	2b	1	1	1
27	J11	1	1	2a	2b	2b	2b	1	2b	1	1	1	2b	2b	2b
28	J12	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	K1	1	1	1	1	5	2b	1	2b	1	1	1	1	2b	2b
30	K2	1	1	1	1	2b	1	1	1	1	1	1	1	1	1
31	K3	1	1	1	1	9	1	1	1	1	1	10	10	2b	9
32	K4	1	1	2b	2b, 5	1	1	1	2b	5, 7	1	2b	2b	2b	2b, 7
33	K5	1	5	2b	2b, 5	2b	2b	2b	2b	2b	2b	10	2b	2b	2b
34	K6	1	2a	1	7	2b	2b	10	1	5, 7	2b	2b	10	3, 5	5
35	K9	1	1	1	1	2b	1	1	2b	1	1	1	1	1	1
36	V1a	1	1	2a	10	2b	10	1	2b	1	3	10	2b	2b	2b
37	V2	1	1	2b	2b	2b	2b	2b	2b	1	2b	1	1**	2b	2b
38	V3	1	1	1	1	2b	2b	1	1	1	1	1**	1**	1	1
39	V4	1	1	1	2b	2b	2b	2b	2b	1sp	2b, 6	2b	2b	2b	1
40	V5	1	1	2a	2b	2b	2b	1	2b	1	1	2b	2b	2b	2b
41	V6	1	1	2a	1	2b	1	1	1	1	1	2b	1	2b	1
42	T1	1	1	2b, 9	2b	2b	2b	2b, 6	2b, 6	10	5, 7	2b, 6	2b, 6	2b, 6	2b, 6
43	T2	1	1	1	1	2b	2b	1	2b	1	5, 7	2b	2b	2b	2b, 3
44	T3	1	1	1	2b	10	2b	7	10	10	5, 7	10	2b	2b	7

45	T4	1	1	1	2b	1	2b	1	2b	1	5, 7	2b	2b	2b	1
46	T5	1	1	1	2b	2b	2b	2b	2b	1	3	2b	2b, 3	1	2b, 3
47	T6	1	1	10	2b	10	10	1	2b	10	10	10	10	10	10
48	T7	1	1	1	2b	5, 7	2b	1	2b	1	1	2b	2b	2b	2b
49	In1	1	1	2b, 5	2b	2b	2b	1	5	2b	10	2b	10	2b	2b
50	In2	1	1	1	2b, 5	2b	2b	2b	2b	2b, 4	5, 7	2b	2b, 5	1	5, 7
51	In3	1	1	2b	2b	2b	2b	2b	2b	1	1	2b	2b	2b	2b
52	S1a	1	2a, 3	10	2b	5	2b	2b	2b	2b	5, 7	10	2b, 5	2b	10
53	S1b	1	5	5	2b, 5	2b	2b	2b	2b, 5	2b	5, 7	2b	2b	2b	2b
54	S2	1	1	2b	10	2b	2b	2b	2b	10	2b	2b	2b	2b	10
55	It1	1	5	2a	1	10	2b	1	2b	10	10	2b	2b	2b	2b
56	B1	1	1	2b	2b	2b	10	1	2b	10	3	3	1	2b	2b

* Subjects who participated in both rounds of data collection with *a* referring to the first round and *b* to the second round.

** Relativisation on GEN subject in lieu of GEN object.

**Appendix VII Individual Performances in the
Grammaticality Judgment Task (Items Relating to Relative Clauses)**

No	Code Name	Relativised positions					Types of non-standard forms**				
		SU	DO		IO	GEN	1	2	3	4	5
		S1	S6	S19	S5	S2	S3	S4	S7	S8	S11
1	M1	2***	1	1	2	1	1	2	2	1	1
2	M2a*	2	1	1	2	1	2	2	2	2	2
3	M2b	1	1	2	1	1	2	2	2	2	2
4	M3	1	2	1	2	2	1	2	1	1	1
5	M4	1	1	1	1	1	1	2	2	2	2
6	M5a	2	1	2	2	1	3	2	2	2	2
7	M5b	1	1	1	2	1	2	2	1	1	3
8	M6	1	1	1	2	1	2	2	2	1	2
9	M7a	1	2	1	2	1	1	2	3	1	3
10	M7b	1	1	1	3	1	2	2	1	1	3
11	M8	2	2	1	2	1	1	2	2	2	2
12	M9	1	1	1	2	1	2	2	2	2	2
13	M10	2	1	1	2	1	1	2	2	2	2
14	M11	1	2	1	2	2	1	2	1	2	2
15	M12	1	1	1	2	1	2	2	2	1	2
16	M13	1	1	1	1	1	2	2	2	2	2
17	M14	2	1	1	1	2	1	1	1	2	2
18	M15	2	1	1	1	1	2	2	2	2	2
19	J1	3	2	2	2	3	3	2	2	1	2
20	J2	1	1	1	1	1	2	2	1	1	2
21	J3	1	1	1	2	2	1	2	2	2	2
22	J4	1	2	1	2	1	2	2	2	2	2
23	J5	1	2	1	2	1	2	2	1	2	1
24	J8	1	1	1	1	1	2	2	1	2	2
25	J9	1	1	1	1	1	2	2	2	2	2
26	J10	1	3	1	1	1	2	2	2	2	2
27	J11	1	1	1	1	1	1	2	1	2	2
28	J12	2	2	1	2	1	2	2	1	2	2
29	K1	1	2	1	2	1	2	2	1	2	2
30	K2	1	2	1	2	1	2	2	2	2	3
31	K3	1	1	1	1	1	2	2	1	2	2
32	K4	1	1	1	2	1	1	2	2	1	2
33	K5	2	1	2	2	2	2	2	2	1	2
34	K6	1	1	1	1	1	2	2	2	1	1
35	K9	1	1	2	1	1	2	1	2	1	2
36	V1a	2	2	2	1	1	2	2	2	1	2
37	V2	2	2	2	1	1	1	2	2	1	2
38	V3	1	2	2	1	1	2	2	2	2	2
39	V4	1	1	2	1	1	2	2	2	2	2
40	V5	1	2	1	2	1	2	2	2	2	2
41	V6	1	1	1	1	1	1	2	2	2	2
42	T1	1	2	2	1	1	3	2	2	3	2
43	T2	1	2	1	2	1	1	2	1	2	2
44	T3	1	3	1	3	2	2	2	1	2	2
45	T4	1	1	1	1	1	1	2	2	2	2
46	T5	2	1	1	1	2	1	2	2	3	2

47	T6	2	1	2	1	1	2	2	2	2	2
48	T7	1	3	2	1	1	2	2	1	2	2
49	In1	1	2	1	1	---	1	2	2	2	2
50	In2	2	1	2	2	2	2	2	2	2	2
51	In3	1	2	2	1	1	2	2	2	2	2
52	S1a	2	2	2	1	1	1	2	2	3	2
53	S1b	2	3	3	2	1	1	1	1	1	3
54	S2	1	2	1	2	3	1	2	2	1	2
55	It1	1	1	1	2	1	1	1	2	2	2
56	B1	2	1	1	2	1	1	2	1	1	2

* Subjects who participated in both rounds of data collection with *a* referring to the first round and *b* to the second round.

** Types of non-standard forms: 1. use of resumptive pronoun; 2. incorrect RC marker – animacy; 3. omission of obligatory RC marker; 4. no adjacency to head noun; and 5. order of Rel N – preposed RC.

*** Coded judgment choices: 1 = acceptable English; 2 = unacceptable English; 3 = not sure

Appendix VIII Counts of Relatives Clauses in the Subjects' Speech and Written Essays

No	Code Name*	The Conversation				Written Essays			Total
		SU	DO	OBL	GEN	SU	DO	OBL	
1	M2n			1		1+ 1*****	1		4
2	M2b					2+ 1*****	1		4
3	M3					1*****			1
4	M4					1+ 2*****	1		4
5	M5a	1				1			2
6	M6					3			3
7	M7a					4	1		5
8	M7b	1				2	1		4
9	M8		1			4		1	6
10	M10		1			1*****	1		3
11	M11					3	1		4
12	M12				1**				1
13	M13	1				1+1***			3
14	M15					2	1		3
15	J2			1**		2			3
16	J8		1	1**/ ****		1	3		6
17	J11		1			1			2
18	JX					1			1
19	K1					1			1
20	K7		1						1
21	K9					1*****			1
22	KX					1			1
23	V1a					3			3
24	V2		1						1
25	V4					4			4
26	V5		1			2			3
27	V6		1****			6+1****	2		10
28	T1					2			2
29	T2	1*****				3			4
30	T3	1							1
31	T4					1	2		3
32	T5	1				3			4
33	T7	2							2
34	In1					1		1**	2
35	In2					3			3
36	In3		2			5+1****+ 1*****			9
37	Stb	1*****				1	2		4
38	F1	3							3
39	B1	1				2+1***	1		5
Total		13	10	3	1	79	18	2	126

* Code names ending in small letters refer to those who participated in both rounds of data collection. Code names ending in X refer to those who dropped out from all the tasks but this one.

** RCs which involve *incorrect RC marker*.

*** RCs which involve no adjacency to the head noun.

**** RCs which involve relativisation on the topic component in a topic-comment structure.

***** RCs which involve *omission of obligatory RC marker*.

**Appendix IX Individual Performances in the Grammaticality
Judgment Task (Items Relating to Implicational Universals) (A)**

No	Code name	Clause Embedding Hierarchy					Subordinate Gap/No Gap Hierarchy		Hierarchy for Complementisers			
		S25	S29	S33	S22	S28	S24	S31	S10	S13	S16	S20
1	M1	1**	1	1	1	1	1	1	2	1	1	1
2	M2a*	2	2	2	1	2	2	3	3	3	2	1
3	M2b	1	2	2	2	2	2	1	2	3	3	1
4	M3	1	1	1	1	1	2	2	3	1	2	3
5	M4	1	1	2	1	1	1	2	1	2	1	2
6	M5a	1	3	2	2	2	3	1	2	1	2	1
7	M5b	1	1	1	3	1	1	1	2	2	1	1
8	M6	2	2	2	1	1	2	2	1	1	1	1
9	M7a	1	1	1	3	1	2	2	3	1	3	2
10	M7b	1	3	3	2	1	1	1	3	3	1	1
11	M8	1	2	1	1	2	2	1	2	1	1	2
12	M9	1	1	1	2	2	2	1	2	2	1	2
13	M10	1	2	2	1	2	2	1	2	2	2	2
14	M11	1	2	2	2	2	2	2	2	2	2	2
15	M12	1	1	1	2	1	2	2	2	2	1	1
16	M13	1	2	2	2	1	2	1	2	1	2	3
17	M14	1	2	2	2	3	2	1	2	2	1	3
18	M15	1	1	2	2	1	2	2	2	2	2	2
19	J1	2	2	2	1	1	1	1	1	1	1	1
20	J2	1	1	2	2	2	1	2	2	1	2	1
21	J3	2	2	1	1	1	2	1	1	1	1	2
22	J4	1	1	2	3	2	2	3	1	2	2	2
23	J5	1	1	2	1	1	1	1	1	1	3	1
24	J8	1	1	3	1	1	2	1	2	3	2	1
25	J9	1	1	2	2	2	1	2	2	2	1	1
26	J10	1	1	2	2	2	2	2	2	2	2	1
27	J11	1	1	2	2	2	2	2	2	1	2	1
28	J12	1	1	2	2	1	1	1	2	2	1	1
29	K1	1	2	2	1	1	2	2	2	2	1	1
30	K2	1	1	2	2	3	2	2	2	2	2	3
31	K3	1	2	1	1	3	2	1	3	1	2	3
32	K4	1	1	1	1	1	3	1	1	1	1	2
33	K5	2	2	1	3	1	2	2	3	2	3	3
34	K6	1	1	2	2	1	2	2	2	2	2	1
35	K9	1	1	1	2	2	2	2	2	2	2	1
36	V1a	1	2	1	2	1	2	1	2	2	2	1
37	V2	1	1	2	1	2	2	2	1	1	1	1
38	V3	1	3	2	1	1	2	1	2	2	2	2
39	V4	2	2	2	3	1	2	1	3	2	1	1
40	V5	1	2	3	1	1	2	1	3	1	1	1
41	V6	1	2	1	1	2	2	1	3	3	1	1
42	T1	2	2	2	1	2	2	3	3	--	1	3
43	T2	1	2	3	1	3	3	3	--	1	3	3
44	T3	3	2	2	1	3	2	2	2	2	3	2
45	T4	1	2	2	2	2	2	2	2	1	1	2

46	T5	1	2	1	1	1	2	1	2	1	2	1
47	T6	1	2	2	3	3	1	2	2	2	1	1
48	T7	2	2	1	2	1	2	1	1	3	1	3
49	In1	3	1	1	2	2	2	1	1	1	1	1
50	In2	1	2	3	1	1	2	2	2	1	1	1
51	In3	2	2	2	2	2	2	2	2	1	2	1
52	S1a	1	2	2	2	2	2	2	3	3	3	3
53	S1b	2	1	2	1	2	2	2	1	2	2	3
54	S2	2	1	1	3	3	1	1	3	1	1	1
55	It1	1	2	1	3	2	2	2	2	1	2	1
56	B1	1	2	2	2	2	2	2	2	2	2	3

* Subjects who participated in both rounds of data collection with *a* referring to the first round and *b* to the second round.

** Coded judgment choices: 1 = acceptable English; 2 = unacceptable English; 3 = not sure

Individual Performances in the Grammaticality Judgment Task (Items Relating to Implicational Universals) (B)

No	Code name	Bridging Verb Hierarchy			Head Noun Phrase Hierarchy			Valency Completeness Hypothesis					
		S18	S21	S26	S14	S17	S23	S27	S34	S30	S35	S32	S36
1	M1	1**	2	1	2	2	1	2	1	2	1	2	1
2	M2a*	1	2	3	2	1	2	2	2	2	1	2	1
3	M2b	1	1	3	2	3	2	2	3	2	1	2	1
4	M3	1	1	1	1	1	1	1	2	1	1	2	1
5	M4	2	1	2	2	2	1	2	2	1	2	1	2
6	M5a	2	3	2	3	2	1	2	3	2	1	2	2
7	M5b	1	3	3	1	2	3	3	3	1	1	3	1
8	M6	2	2	2	2	2	2	3	2	2	1	2	1
9	M7a	2	3	3	3	1	2	1	1	3	2	2	1
10	M7b	1	3	3	1	3	3	3	3	2	1	2	1
11	M8	1	2	2	2	1	1	2	1	2	1	2	2
12	M9	1	1	2	1	1	1	3	3	2	1	2	2
13	M10	1	2	2	2	2	2	2	2	2	1	2	1
14	M11	2	2	2	2	2	2	2	2	2	1	2	2
15	M12	1	1	1	2	1	2	1	2	2	1	1	1
16	M13	2	2	1	3	3	1	2	1	1	1	2	2
17	M14	1	1	2	2	2	2	1	1	2	2	2	1
18	M15	2	2	2	2	1	2	2	2	2	2	2	1
19	J1	2	2	2	3	1	1	2	1	1	1	1	1
20	J2	1	1	2	2	2	2	2	2	2	2	2	2
21	J3	1	1	1	2	1	1	1	2	2	1	2	1
22	J4	2	2	2	1	1	3	2	2	2	3	2	3
23	J5	2	2	3	1	1	1	2	2	2	1	2	1
24	J8	1	2	2	2	1	1	3	2	2	3	2	1
25	J9	2	2	2	1	1	1	2	2	2	1	2	1
26	J10	2	2	2	2	1	2	2	2	2	1	2	2
27	J11	2	2	2	1	1	2	2	2	2	1	2	2
28	J12	2	2	2	2	2	2	2	2	2	1	2	1
29	K1	2	2	1	1	1	1	2	1	1	1	2	1

30	K2	2	3	2	1	2	2	2	1	2	1	2	2
31	K3	2	2	2	2	2	2	2	2	2	3	2	3
32	K4	2	2	2	2	2	2	1	2	1	2	2	2
33	K5	1	1	1	1	1	1	1	2	3	1	2	1
34	K6	1	2	3	3	1	1	1	1	1	1	2	1
35	K9	2	2	1	1	1	1	2	3	3	2	2	1
36	V1a	2	2	3	1	3	2	2	2	1	1	2	2
37	V2	2	1	1	2	1	1	2	1	2	1	2	2
38	V3	1	1	1	1	1	2	2	2	2	1	2	2
39	V4	1	1	1	2	3	2	2	1	2	1	2	2
40	V5	1	2	3	2	2	2	2	2	2	1	2	1
41	V6	2	1	1	3	1	2	3	2	2	1	2	1
42	T1	3	3	3	3	3	3	2	2	2	3	2	2
43	T2	1	3	3	3	1	3	3	3	3	3	2	1
44	T3	2	2	2	1	1	3	3	3	1	1	2	1
45	T4	1	1	1	2	2	1	2	1	1	2	2	2
46	T5	2	2	2	1	1	1	3	2	1	3	1	2
47	T6	2	2	2	1	1	2	1	1	1	2	1	1
48	T7	1	3	3	3	1	2	2	2	2	1	2	1
49	In1	2	3	2	2	1	2	2	2	2	2	2	2
50	In2	2	3	3	1	1	1	3	2	1	2	1	3
51	In3	2	2	2	2	2	2	2	2	2	2	2	1
52	S1a	2	3	2	2	3	1	3	2	2	2	2	2
53	S1b	2	3	3	3	1	2	2	2	2	1	2	1
54	S2	1	3	3	2	2	2	3	1	2	2	1	1
55	It1	1	2	2	2	2	2	2	2	2	2	2	1
56	B1	2	2	2	2	2	1	2	3	1	1	2	1

* Subjects who participated in both rounds of data collection with *a* referring to the first round and *b* to the second round.

** Coded judgment choices: 1 = acceptable English; 2 = unacceptable English; 3 = not sure

**Appendix X Aggregate Counts of Instances
of Different Features in the Conversation Data**

No	Code Name	Total Words Spoken	Features**								
			1	2	3	4***	5	6	7	8	9
1	M1	929	2	2	6		10	1	1	4	
2	M2a*	706	3		4	3 (1)	6				1
3	M2b	1268	6		8	2 (1)	9	1		3	1
4	M3	804	3	1	4	1	15	1	4	6	2
5	M4	762	3		10	3	3	1		2	
6	M5a	851	1	1	3		7		2	6	2
7	M5b	829	3		3	(2)	6		2	4	2
8	M6	870	1	2	12	2	15		1	3	2
9	M7a	899	4		23	1	11		4	6	2
10	M7b	797	6	3	19	2	6	1		5	2
11	M8	731	2	1	12		10	2	2		1
12	M9	549	5	2	7		3	1	3		2
13	M10	1164	9	3	6	2	16	1	3	6	1
14	M11	585	5	1	2	1	7			1	
15	M12	952	4		11	1	8	1	3	2	5
16	M13	1129	7		14	2	11	1	4	2	3
17	M14	637	4		15		4	1	4	1	2
18	M15	592	3	4	7	(1, 20)	4	2	1	1	
19	J1	581			7		8	1		1	
20	J2	605	3		10	2 (1)	2		1	1	2
21	J3	802	6		17		15		1	1	1
22	J4	493	4		9	1	3	1	2	3	
23	J5	778			13	(1)	9	4		5	2
24	J6	404	3		7		2		6		6
25	J7	154			1		2				1
26	J8	1239	17	1	22	6	17	2	1	4	5
27	J9	497			7		3			2	1
28	J10	1079	1	1	18		5		3	4	3
29	J11	531	2		10		4			3	6
30	K1	1742	10	5	18		21	2	3	3	3
31	K2	700	4	1	11		2			1	3
32	K3	704	4	3	6	1	1		4	2	2
33	K4	363	4		2		5			5	
34	K5	656	2	1	11		2		6	2	2
35	K6	739	3		9	(1)	8			1	2
36	K7	664	5		13	1	3	1	4	2	5
37	K8	841	6	1	16	2 (1)	6		4	2	5
38	V1a	734	4	2	6	(7)	7		4	2	
39	V2	805	2		13	(2)	11	1	3	1	5
40	V3	632			2		10		8	11	1
41	V4	1023	6		14	2	6	1	3	4	2
42	V5	1221	7	1	12	2	10		4	3	2
43	V6	1122	3	1	8	1	16	5	6	5	5
44	V7	929	4	1	9	3	8		9	1	5
45	T1	747	2		7		6		1	3	2
46	T2	665	5		4		14		3	4	3
47	T3	710	5		15	1	4		4	7	1
48	T4	520	4	1	6	1	4	2	4	3	
49	T5	607	3	2	6		10		2	1	2
50	T6	1003	5		11	1 (1)	8	1	3	4	1
51	T7	632	4		5	(1)	6			4	
52	In1	550	3		5	1 (1)	9		1	1	

53	In2	915	3		8	2	9		1	5	1
54	In3	865	9		6	2 (3)	2		2	2	2
55	S1a	1071	3		2	(2)	4		8	8	2
56	S1b	986	5		8	(2,1i)	5	1	3	4	6
57	S2	686	4	1	8	(1)	8		3	4	3
58	F1	1081	10	1	25	5	6			1	
59	It1	770	6	1	17	3 (1,1i)	10	1	1	1	
60	B1	652	14	3	6	3	4		1	7	
Total		47552	256	47	576	60/ (34)	446	37	143	180	120

* Subjects who participated in both rounds of data collection with *a* referring to the first round and *b* to the second round.

** Coded features are as follows: 1 = filler-gap construction; 2 = passive voice; 3 = structure of sentential coordination; 4 = structure of sentential subordination (finite complement clauses); 5 = structure of sentential subordination (finite adverbial clauses); 6 = topic-comment structure; 7 = non-standard use of subject-verb agreement; 8 = non-standard use of noun number; and 9 = omission of required indefinite articles.

*** Figures within parentheses and outside refer respectively to finite complement clauses containing *that*, *if* and *whether* and other types of finite complement clauses.

Appendix XI Aggregate Counts of Instances of Different Features in the Subjects' Written Essays

No	Code Name*	Total Words Written	Features**								
			1	2	3	4***	5	6	7	8	9
1	M1	82			1		2		2	1	1
2	M2a	352	4	1	6	4 (3)	8		2		3
3	M2b	470	7	2	9	6 (5)	9	1	2	1	
4	M3	88	2			(2)	3			1	
5	M4	248	5		5	1 (6)	2			2	
6	M5a	238	2		1	(5)	6			1	
7	M6	425	5	1	5	4 (4)	9		3	4	2
8	M7a	296	5	3	3		4	2			
9	M7b	295	3	4	6		6	1	1	1	
10	M8	498	5	3	10	4	8				
11	M10	170	2	1	6	(1)			4	3	
12	M11	266	5	3	5	1	2		5	9	
13	M12	155			2	(1)					
14	M13	365	2	3	7	(3)	3		1	1	1
15	M15	154	3	3	3	1	2			4	1
16	J1	99			1		1		1	1	
17	J2	149	2	2		1	3		1		
18	J3	91			4		2		1	1	
19	J8	373	7	1	5	5 (1)	8	2		1	
20	J11	146	2	5	3	7 (1)	2			5	1
21	JX	55	1		1					1	
22	K1	193	2		2	(1)	6		1	1	
23	K3	290	1	2	4	1 (3)	5			4	
24	K5	176		8	5	2			4	6	1
25	K9	96	2	3	3	1	1		1	2	
26	KX	118	1		4	(1)					1
27	V1a	246	4	1	3					1	
28	V2	374	1		11	4	5	2		2	
29	V4	284	4	2	5	1 (2)			2	2	
30	V5	136	2	11	1				1		
31	V6	473	10	2	6	6 (3)	8	2			1
32	T1	157	5	4	2	3 (2)	3		1	3	1
33	T2	348	3	5	1		4				1
34	T4	300	5	2	4	1 (3)	5		2	1	1
35	T5	201	4		3	1 (2)	3		1	7	
36	T7	226			4	(1)	1			7	
37	In1	261	3	2	10	2	2	4	1	1	2
38	In2	276	4	1	2		2				
39	In3	275	9	2	3	1 (4)	8				
40	S1a	139								2	
41	S1b	310	3	3	7	(3, 1w)	9	2		1	
42	S2	107		1	1	(4)	2			1	1
43	It1	163		2	8	(1)		2			
44	B1	270	4	10	2	1 (1)	3			1	1
45	MaX	141	1		3	1	5				
Total		10575	130	93	177	59/ (64)	152	18	37	79	19

* Code names ending in small letters refer to those who participated in both rounds of data collection. Code names ending in X refer to those who dropped out from all the tasks but this one. (MaX: a subject speaking Malay).

** Coded features are as follows: 1 = filler-gap construction; 2 = passive voice; 3 = structure of sentential coordination; 4 = structure of sentential subordination (finite complement clauses); 5 = structure of sentential subordination (finite adverbial clauses); 6 = topic-comment structure; 7 = non-standard use of subject-verb agreement; 8 = non-standard use of noun number; and 9 = omission of required indefinite articles.

*** Figures within parentheses and outside refer respectively to finite complement clauses containing *that*, *if* and *whether* and other types of finite complement clauses.

Bibliography

- Adjémian, C. and Liceras, J. (1984). Accounting for adult acquisition of relative clauses: Universal grammar, L1, and structuring the intake. In F. Eckman, L. Bell, and D. Nelson (Eds.), *Universals of Second Language Acquisition* (pp. 101-18). Rowley, MA: Newbury House.
- Andersen, R. W. (1978). An implicational model for second language research. *Language Learning* 28, 221-82.
- Aarts, F., and Schils, E. (1995). Relative clauses, the Accessibility Hierarchy and the Contrastive Analysis Hypothesis. *IRAL* 33, 47-63.
- Bailey, C. J. (1971). Trying to talk in the new paradigm. *University of Hawaii Working Papers in Linguistics* 4, 312-39.
- Bailey, N., Madden, C., and Krashen, S. (1974). Is there a "natural sequence" in adult second language learning?. *Language Learning* 24, 235-43.
- Bates, E., and MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner, and L. Gleitman (Eds.), *Language Acquisition: The State of the Art* (pp. 173-218). New York: Cambridge University Press.
- Bayley, R. (1996). Competing constraints on variation in the speech of adult Chinese learners of English. In R. Bayley, and D. R. Preston (Eds.), *Second Language Acquisition and Linguistic Variation* (pp. 97-120). Amsterdam: John Benjamins.
- Berlin, B., and Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.
- Bickerton, D. (1971). Inherent variability and variable rules. *Foundations of Language* 7, 457-92.
- Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. Gass, and J. Schachter (Eds.), *Linguistic Perspectives on Second Language Acquisition* (pp. 41-68). Cambridge: Cambridge University Press.
- Bowe, H. (1990). *Categories, Constituents, and Constituent Order in Pitjantjatjara: An Aboriginal Language of Australia*. London and New York: Routledge.
- Brines, V. (1990). Learning yes-no questions under different academic conditions. *IRAL* 28, 234-47.

- Brooks, N. (1964). *Language and Language Learning: Theory and Practice* (2nd edition). New York: Harcourt Brace and World.
- Brown, R. (1973). *A First Language*. Cambridge, MA: Harvard University Press.
- Bybee, J. L. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.
- Carlisle, R. S. (1997). The modification of onsets in a markedness relationship: Testing the interlanguage structural conformity hypothesis. *Language Learning* 47, 327-61.
- Chen, S. (1990). A study of communication strategies in interlanguage production by Chinese EFL learners. *Language Learning* 40, 155-87.
- Chomsky, N. (1959). Review of "Verbal Behavior" by B. F. Skinner. *Language* 35, 26-58.
- (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- (1973). Conditions on transformations. In S. R. Anderson and P. Kiparsky (Eds.), *Festschrift for Morris Halle* (pp. 232-86). New York: Holt, Rinehart and Winston.
- (1981a). Principles and parameters in syntactic theory. In N. Hornstein, and D. Lightfoot (Eds.), *Explanation in Linguistics* (pp. 32-75). London: Longman.
- (1981b). *Lectures on Government and Binding*. Dordrecht: Foris.
- (1986). *Barriers*. Cambridge, MA: MIT Press.
- (1993). A minimalist program for linguistic theory. In K. Hale, and S. J. Keyser (Eds.), *The View from Building 20* (pp. 1-52). Cambridge, MA: MIT Press.
- Clahsen, H. (1988). Parameterized grammatical theory and language acquisition: A study of the acquisition of verb placement and inflection by children and adults. In S. Flynn, and W. O'Neil (Eds.), *Linguistic Theory in Second Language Acquisition* (pp. 47-75). Dordrecht: Kluwer.
- Comrie, B. (1978). Ergativity. In W. Lehmann (Ed.), *Syntactic Typology* (pp. 329-94). Austin: University of Texas Press.
- (1981). *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell Publisher Limited.

- (1984). Why linguists need language acquirers. In W. E. Rutherford (Ed.), *Language Universals and Second Language Acquisition* (pp. 11-32). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Cook, V. (1973). The comparison of language development in native children and foreign adults. *IRAL* 11, 13-28.
- (1977). Cognitive processes in second language learning. *IRAL* 15, 1-20.
- Corder, S. P. (1967). The significance of learners' errors. *IRAL* 2-3, 161-70.
- Croft, W. (1990). *Typology and Universals*. Cambridge: Cambridge University Press.
- (1995). Modern syntactic typology. In M. Shibatani, and T. Bynon. (Eds.), *Approaches to Language Typology* (pp. 85-144). New York: Clarendon Press.
- Denning, K., and Kemmer, S. (Eds.) (1990). *On Language: Selected Writings of Joseph H. Greenberg*. Stanford: Stanford University Press.
- Dixon, R. M. W. (1979). Ergativity. *Language* 55, 59-138.
- Döpke, S. (1998). Competing language structures: The acquisition of verb placement by bilingual German-English children. *Child Language* 25, 555-84.
- Dryer, M. S. (1986). Primary objects, secondary objects and anti-dative. *Language* 62, 808-45.
- (1988). Object-verb order and adjective-noun order: Dispelling a myth. *Lingua* 74, 185-217.
- (1989). Large linguistic areas and language sampling. *Studies in Language* 13, 257-92.
- (1991). SVO languages and the OV:VO typology. *Journal of Linguistics* 27, 443-82.
- (1992). The Greenbergian word order correlations. *Language* 68, 81-138.
- Dulay, H., and Burt, M. (1973). Should we teach children syntax?. *Language Learning* 23, 245-58.
- (1974a). Errors and strategies in child second language acquisition. *TESOL* 8, 129-36.
- (1974b). Natural sequences in child second language acquisition. *Language Learning* 24, 37-53.
- (1974c). A new perspective on the creative constructive process in child second language acquisition. *Language Learning* 24, 253-78.

- Eckman, F. (1977). Markedness and the Contrastive Analysis Hypothesis. *Language Learning* 27, 315-30.
- (1991). The Structural Conformity Hypothesis and the Acquisition of Consonant Clusters in the Interlanguage of ESL Learners. *Studies in Second Language Acquisition* 13, 23-41.
- (1993). On the natural domain of linguistic universals. In M. Eid, and G. Iverson (Eds.), *Principles and Prediction: The Analysis of Natural Language: Papers in Honor of Gerald Sanders* (pp. 45-69). Amsterdam & Philadelphia: John Benjamins Publishing Co.
- (1996). A functional-typological approach to second language acquisition theory. In W. Ritchie, and T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 195-211). San Diego: Academic Press, Inc.
- Eckman, F., Bell, L., and Nelson, D. (1988). On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics* 9, 1-20.
- Eckman, F., Moravcsik, E., and Wirth, J. (1989). Implicational universals and interrogative structures in the interlanguage of ESL learners. *Language Learning* 39, 173-205.
- Ellis, R. (1985). Sources of variability in interlanguage. *Applied Linguistics* 6, 118-31.
- (1992). *Second Language Acquisition and Pedagogy*. Clevedon: Multilingual Matters.
- (1994). *The Study of Second Language Acquisition*. New York: Oxford University Press.
- Epstein, S. D., Flynn, S., and Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences* 19, 677-714.
- (1998). The strong continuity hypothesis: Some evidence concerning functional categories in adult L2 acquisition. In S. Flynn, G. Martohardjono, and W. O'Neil. (Eds.), *The Generative Study of Second Language Acquisition* (pp. 61-77). Mahwah, NJ: Lawrence Erlbaum Associates.

- Flynn, S. (1984). A universal in L2 acquisition based on a PBD typology. In F. Eckman, L. Bell, and D. Nelson (Eds.), *Universals of Second Language Acquisition* (pp. 75-87). Rowley, MA: Newbury House.
- (1987). Contrast and construction in a parameter-setting in L2 acquisition. *Language Learning* 37, 19-62.
- (1989). The role of the head-initial/head-final parameter in the acquisition of English relative clauses by adult Spanish and Japanese speakers. In S. Gass, and J. Schachter (Eds.), *Linguistic Perspectives on Second Language Acquisition* (pp. 89-108). Cambridge: Cambridge University Press.
- (1996). A parameter-setting approach to second language acquisition. In W. Ritchie, and T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 121-58). San Diego: Academic Press, Inc.
- Fox, B. A. (1987). The noun phrase accessibility hierarchy reinterpreted: Subject primacy or the absolutive hypothesis? *Language* 63, 856-70.
- Gass, S. (1979). Language transfer and universal grammatical relations. *Language Learning* 29, 327-44.
- (1982). From theory to practice. In M. Hines, and W. Rutherford (Eds.), *On TESOL '81*. Washington, DC: TESOL.
- (1989). Language universals and second-language acquisition. *Language Learning* 39, 497-534.
- (1994). The reliability of second-language grammaticality judgments. In E. Tarone, S. Gass, and A. D. Cohen (Eds.), *Research Methodology in Second-Language Acquisition* (pp. 303-22). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gass, S., and Ard, J. (1980). L2 data: Their relevance for language universals. *TESOL Quarterly* 14, 443-52.
- Givón, T. (1979). *On Understanding Grammar*. New York: Academic Press.
- (1980). The binding hierarchy and the typology of complements. *Studies in Language* 4, 333-77.
- (1984). *Syntax: A Functional-Typological Introduction*. Vol. I. Amsterdam and Philadelphia: John Benjamins.

- Greenberg, J. (1954). A quantitative approach to the morphological typology of language. In R. Spencer (Ed.), *Methods and Perspectives in Anthropology?* (pp. 192-220). Minneapolis.
- (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (Ed.), *Universals of Language* (pp. 73-113). Cambridge, MA: MIT Press.
- (1966). Language universals, with special reference to feature hierarchies. (*Junua Linguarum*, Series Minor 59). The Hague: Mouton.
- 1969. Some methods of dynamic comparison in linguistics. In K. Denning, and S. Kemmer (Eds.), (1990). *On Language: Selected Writings of Joseph H. Greenberg*. Stanford: Stanford University Press.
- (1976). *Language Universals*. The Hague: Mouton.
- (1979). Rethinking linguistics diachronically. *Language* 55, 275-90.
- (1991a). Typology/universals and second language acquisition. In T. Huebner, and C. A. Ferguson (Eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories*. US: John Benjamins B. V.
- (1991b). Two approaches to language universals. In L. R. Waugh, and S. Rudy (Eds.), *New Vistas in Grammar: Invariance and Variation* (pp. 417-35). Amsterdam: John Benjamins Publishing Company.
- Greenberg, J. H., Charles A. F., and Moravcsik, E. A. (Eds.) (1978). *Universals of Human Language* (4 vols.). Stanford: Stanford University Press.
- Gregg, K. (1984). Krashen's monitor and Occam's razor. *Applied Linguistics* 5, 79-100.
- Haiman, J. (1980). The iconicity of grammar: isomorphism and motivation. *Language* 56, 515-40.
- (1985). *Natural Syntax*. Cambridge: Cambridge University Press.
- Hatch, E., and Lazaraton, A. (1991). *The Research Manual: Design and Statistics for Applied Linguistics*. Boston, MA: Heinle & Heinle Publishers.
- Hawkins, J. A. (1983). *Word Order Universals*. New York: Academic Press, Inc.
- (1985). The suffixing preference: A processing explanation. *Linguistics* 23, 723-58.
- (1987). Implicational universals as predictors of language acquisition. *Linguistics* 25, 453-73.

- (1988a). Prefixing and suffixing universals in relation to basic word order. *Lingua* 74, 219-59.
- (1988b). On generative and typological approaches to universal grammar. *Lingua* 74, 85-100.
- (1990a). Seeking motives for change in typological variation. In *Studies in Typology and Diachrony: Papers Presented to Joseph H. Greenberg on his 75th Birthday* (pp. 95-128). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- (1990b). A parsing theory of word order universals. *Linguistic Inquiry* 21, 223-61.
- (1993). Heads, parsing and word-order universals. In G. G. Corbett, N. M. Fraser, and S. McGlashan (Eds.), *Heads in Grammatical Theory* (pp. 231-265). Cambridge: Cambridge University Press.
- (1994). *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- (1997). Some issues in a performance theory of word order. In A. Siewierska (Ed.), *Constituent Order in the Languages of Europe* (pp. 729-81). Berlin: Mouton de Gruyter.
- (1999). Processing complexity and filler-gap dependencies across grammars. *Language* 75, 244-85.
- (2001). Why are categories adjacent? *Journal of Linguistics* 37, 1-34.
- Hawkins, J. A., and Gilligan, G. (1988). Prefixing and suffixing universals in relation to basic word order. *Lingua* 74, 219-59.
- Hopper, P., and Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language* 56, 251-99.
- Huebner, T. (1983). *A Longitudinal Analysis of the Acquisition of English*. Ann Arbor: Karoma Publishers.
- Hyltenstam, K. (1984). The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses. In R. Andersen (Ed.), *Second Languages: A Cross-Linguistic Perspective* (pp. 39-58). Rowley, Mass: Newbury House Publishers.

- Ito, A. (2000). Tests as a second language research method: Their types, reliability, validity, and variable research results. *ITL, Review of Applied Linguistics* 127-128, 1-36.
- (2001). Japanese EFL learners' sensitivity to configurational distinction in English relativization. *ITL, Review of Applied Linguistics* 131-132, 11-33.
- Joseph, B. D. (1983). Relativization in modern Greek: Another look at the accessibility hierarchy constraints. *Lingua* 60, 1-24.
- Juffs, A. (1990). Tone, syllabus structure and interlanguage phonology: Chinese learners' stress errors. *IRAL* 28, 99-117.
- Kay, P., and McDaniel, C. K. (1978). The linguistic significance of the meaning of basic color terms. *Language* 54, 610-46.
- Keenan, E. L. (1985). Relative clauses. In T. Shopen (Ed.), *Language Typology and Syntactic Description* 2 (pp. 141-170). Cambridge: Cambridge University Press.
- Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8, 63-99.
- (1979). Data on the noun phrase accessibility hierarchy. *Language* 55, 333-51.
- Kessler, C., and Idar, I. (1979). The acquisition of English by a Vietnamese mother and child. *Working Papers on Bilingualism* 18, 65-77.
- Kim, S. (1999). The subject-object asymmetry in the acquisition of WH-questions by Korean learners of English. *Journal of Pan-Pacific Association of Applied Linguistics* 3, 1-11.
- Krashen, S. (1977). Some issues relating to the Monitor Model. In H. D. Brown et al. (Eds.), *Teaching and Learning English as a Second Language: Trends in Research and Practice, On TESOL '77*. Washington D. C.: TESOL.
- (1980). The input hypothesis. In J. E. Alatis (Ed.), *Current Issues in Bilingual Education*. Washington D. C.: Georgetown University Press.
- (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press.
- Krashen, S., and Terrell, T. (1983). *The Natural Approach: Language Acquisition in the Classroom*. Oxford: Pergamon Press.

- Kluender, R. (1992). Deriving island constraints from principles of prediction. In H. Goodluck, and M. Rochemont (Eds.), *Island Constraints: Theory, Acquisition and Processing* (pp. 223-53). Dordrecht: Kluwer.
- Kumpf, L. (1983). A case study of temporal reference in interlanguage. In C. Campbell et al (Eds.), *Proceedings of the Los Angeles Second Research Forum, Volume II* (pp.179-92). Los Angeles: Dept. of English, ESL Section, UCLA.
- Kupferberg, I., and Olshtain, E. (1996). Explicit contrastive instruction facilitates the acquisition of difficult L2 forms. *Language-Awareness* 5, 149-65.
- Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language* 45, 715-62.
- Lado, R. (1983). Procedures in comparing two grammatical structures. In B. W. Robinett, and J. Schachter (Eds.), *Second Language Learning: Contrastive Analysis, Error Analysis and Related Aspects*. Ann Arbor: The University of Michigan.
- Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult learners of English. Unpub. Ph.D. Dissertation, University of Michigan.
- Larsen, T., and Norman, W. (1979). Correlates of ergativity in Mayan grammar. In F. Plank (Ed.), *Ergativity* (pp.347-70). New York: Academic Press.
- Lawal, N. S. (1987). Yoruba relativisation and the continuous segment principle. *Studies in African Linguistics* 18, 67-79.
- Lehmann, W. P. (1973). A structural principle of language and its implications. *Language* 49, 47-66.
- (1978). The great underlying ground-plans. In W. P. Lehmann (Ed.), *Syntactic Typology: Studies in the Phenomenology of Language*. Austin: University of Texas Press.
- Lehmann, C. (1986). On the typology of relative clauses. *Linguistics* 24, 663-80.
- Lewis, M. (1985). Relative clauses in A(eng)l(open o) Ewe. *Studies in African Linguistics* supplement 9, 196-202.
- Li, C. M., and Li, L. (1994). Relativization in English and Chinese. *Papers and Studies in Contrastive Linguistics* 29, 61-75.
- Mallinson, G., and Blake, B. J. (1981). *Language Typology: Cross-Linguistic Studies in Syntax*. Amsterdam: North-Holland Publishing Company.

- Manaster-Ramer, A. (1979). The other side of accessibility. *Papers from the Regional Meetings. Chicago Linguistic Society* 15, 207-19.
- Mazurkewich, I. (1984). Dative questions and markedness. In F. Eckman, L. Bell, and D. Nelson (Eds.), *Universals of Second Language Acquisition* (pp. 119-31). Rowley, MA: Newbury House.
- Mukattash, L. (1981). Wh-questions in English: A problem for Arab students. *IRAL* 19, 317-32.
- Munnich, E., Flynn, S., and Martohardjono, G. (1994). Elicited imitation and grammaticality judgment tasks: What they measure and how they relate to each other. In E. E. Tarone, S. M. Gass, and A. D. Cohen (Eds.), *Research Methodology in Second-Language Acquisition* (pp. 227-43). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Park, H. S. (1998). The role of the head-directional parameter in second language learning: Especially for Korean learners of English. *Journal of Japan-Korea Association of Applied Linguistics* 1, 33-48.
- (2000). Markedness and learning principles in SLA: Centering on acquisition of relative clauses. *Journal of Pan Pacific Association of Applied Linguistics* 4, 87-114.
- Pavesi, M. (1986). Markedness, discoursal modes, and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition* 8, 38-55.
- Prevost, P., and White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16, 103-33.
- Richards, J. C., Platt, J., and Platt, H. (1992). *Longman Dictionary of Language Teaching and Applied Linguistics* (2nd ed). Essex: Longman Group UK Limited.
- Rosansky, E. (1976). Methods and morphemes in second language acquisition. *Language Learning* 26, 409-25.
- Ross, J. R. (1967). *Constraints on Variables in Syntax*. Cambridge, MA: MIT dissertation.

- Rutherford, W. (1984). Introduction. In W. E. Rutherford (Ed.), *Language Universals and Second Language Acquisition* (pp. 1-9). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Sadighi, F. (1994). The acquisition of English restrictive relative clauses by Chinese, Japanese, and Korean adult native speakers. *IRAL* 32, 141-53.
- Sadighi, F., and Jafarpur, A. (1994). Is there a role for learner treatment in comprehending relative clauses? *RELC Journal* 25, 56-76.
- Sato, C. J. (1990). *The Syntax of Conversation in Interlanguage Development*. Germany: Gunter Narr Verlag Tübingen.
- Saunders, N. J. (1983). "Will you get off this plane somewhere?" Form and function in the L2 acquisition of English by adult Japanese. *Occasional Papers Applied Linguistics Association of Australia* 6, 66-83.
- Schachter, J. (1989). Testing a proposed universal. In S. Gass, and J. Schachter (Eds.), *Linguistic Perspectives on Second Language Acquisition*. Cambridge: Cambridge University Press.
- Schumann, J. (1976). Second language acquisition: The pidginization hypothesis. *Language Learning* 26, 391-408.
- (1978). *The Pidginization Process: A Model for Second Language Acquisition*. Rowley, MA: Newbury House.
- (1986). Research on the acculturation model for second language acquisition. *Journal of Multilingual and Multicultural Development* 7, 379-92.
- (1990). Extending the scope of the acculturation/pidginization model to include cognition. *TESOL Quarterly* 24, 667-84.
- (2001). Appraisal psychology, neurobiology, and language. *Annual Review of Applied Linguistics* 21, 23-42.
- Seiler, H. (1995). Cognitive-conceptual structure and linguistic encoding: Language universals and typology in the UNITYP framework. In M. Shibatani, and T. Bynon. (Eds.), *Approaches to Language Typology* (pp. 273-325). New York: Clarendon Press.

- Seliger, H. W. (1984). Processing universals in second language acquisition. In F. Eckman, L. Bell and D. Nelson. (Eds.), *Universals of Second Language Acquisition*. Rowley: Newbury House Publishers, Inc.
- Seliger, H. W., and Shohamy, E. (1989). *Second Language Research Methods*. Oxford: Oxford University Press.
- Selinker, L. (1972). Interlanguage. *IRAL* 10, 209-31.
- Shibatani, M., and Bynon, T. (Eds...) (1995). *Approaches to Language Typology*. New York: Clarendon Press.
- Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics* 4, 142-64.
- (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning* 35, 373-403.
- (1988). *Variation in Interlanguage*. London: Edward Arnold.
- Towell, R., and Hawkins, R. (1994). *Approaches to Second Language Acquisition*. Clevedon: Multilingual Matters Ltd.
- Tushyeh, H. Y. (1988). Transfer and related strategies in the acquisition of English¹: relative clauses by adult Arab learners. *Papers and Studies in Contrastive Linguistics* 23, 69-86.
- Ullan, R. (1978). Interrogative systems. In Greenberg (Ed.) (1978), Volume 4, 211-48.
- VanPatten, B. (1984). Processing strategies and morpheme acquisition. In F. Eckman, L. Bell, and D. Nelson (Eds.), *Universals of Second Language Acquisition* (pp. 88-98). Rowley, MA: Newbury House.
- Vennemann, T. (1974). Analogy in generative grammar: the origin of word order. In *Proceedings of the Eleventh International Congress of Linguistics*. Bologna: Il Mulino.
- de Villiers, J., and de Villiers, P. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research* 2, 267-78.
- Wei, L. (1997). Variation in the acquisition of morpheme types in the interlanguage of Chinese and Japanese learners of English as a second language. *Dissertation Abstracts International, -A: The Humanities and Social Sciences* 57, 7, 3003-A.

- Whaley, L. J. (1997). *Introduction to Typology: The Unity and Diversity of Language*. Thousand Oaks: SAGE Publications.
- White, L. (1985). The 'Pro-drop' parameter in adult second language acquisition. *Language Learning* 35, 47-62.
- (1987). Against comprehensible input: The input hypothesis and the development of second-language competence. *Applied Linguistics* 8, 95-110.
- (1989). Linguistic universals, markedness and learnability: Comparing two different approaches. *Second Language Research* 5, 127-40.
- White, L., and Juffs, A. (1998). Constraints on wh-movement in two different contexts of nonnative language acquisition: Competence and processing. In S. Flynn, G. Martohardjono, and W. O'Neil. (Eds.), *The Generative Study of Second Language Acquisition* (pp. 111-29). Mahwah, NJ: Lawrence Erlbaum Associates.
- Yeoh, C. K. (1977). Restrictive relative clauses in Bahasa Malaysia. *Studies in the Linguistic Sciences* 7, 82-105.
- Yip, V., and Matthews, S. (1995). I-interlanguage and typology: the case of topic-prominence. In L. Eubank, L. Selinker, and M. Sharwood Smith (Eds.), *The Current State of Interlanguage: Studies in Honor of William E. Rutherford* (pp. 17-30). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Young, R. (1991). *Variation in Interlanguage Morphology*. New York: Peter Lang Publishing Inc.
- Zobl, H. (1995). Markedness aspects of case-marking in L1 French/L2 English. In L. Eubank, L. Selinker, and M. Sharwood Smith (Eds.), *The Current State of Interlanguage: Studies in Honor of William E. Rutherford* (pp. 197-203). Amsterdam/Philadelphia: John Benjamins Publishing Co.