

**Form and Function in the Interlanguage of Zairean Learners  
of English**

**Bisimwa Chilange Ntahwakuderwa**

**Submitted in fulfilment of the requirements  
for the Degree of Doctor of Philosophy in  
the Faculty of Arts of the  
University of Edinburgh, 1987  
Edinburgh, Scotland**



**Dedication**

To my parents with gratitude for so much love, affection and abnegation.

## **DECLARATION**

**I hereby declare that the present thesis has been composed by myself, and that the work included in it is entirely my own.**

## Acknowledgements

First and foremost I wish to thank the authorities of the University of Kinshasa for granting me leave of absence and the British Council for financial support, without which my plans for a Ph.D. would have remained an unattainable goal.

Secondly, throughout my period of Ph.D. research, I benefited from the unfailing moral support and intellectual stimulation given by my supervisor, Dr. Alan Davies, who in spite of his numerous duties was always able not just to lend me a listening ear, but to provide lucid advice and constructive criticism. I unreservedly owe him my life-long gratitude. I also wish to thank the marvellous staff of the Faculty of Arts at the University of Edinburgh for enabling me to benefit from their outstanding professional and academic scholarship. Most notably, my sincere thanks go to the other members of my research committee: Anthony Howatt and Alan Kemp; as well as to Professor Jim Hurford, Dr. Ellen Bard, Keith Mitchell, Christine Nuttall, and Antonella Sorace. I sincerely hope that their encouragement will be extended throughout my professional career.

My efforts in working towards this academic goal would not have been even half as manageable as it turned out to be without the assistance of so many other staff and friends, especially as regards such inevitable chores as computing, administrative paperwork, typing, etc. I would like to record my most sincere thanks to Irene MacLeod and Norman Dryden for their help with the computer; Ann Harrison, my programme officer; Wakilongo Kisali, of the British Embassy in Kinshasa, and Sylvia Motherwell who helped me with the typing.

Many remarkable colleagues and friends, through discussions and their comments, helped me sharpen my own ideas or working hypotheses. Owing to shortage of space I will mention only a few: Professor Vunduawe Te Pemako, Dr. Carew Treffgarne, Dr. Helen Borland, Dr. John Maher, Dr. Maria Pavesi, Dr. Kimani Njoroge, Dr. Chizungu Rudahindwa, Dr. Ndoma Ungina, Dr. Ngangura Kasole, Catherine Wallace, John Landon, Regina Weinert, Sinfree Makoni, Hisham Ali, Rosemary Baker, Gibson Ferguson, Denise Rokosz, Gladys Tang, Nilofar Durrani, Pindi Makaya, Tshimankinda Biselela, Katsuva Ngoloma, and Kamana Tshibengabo.

The most pleasant duty of all is for me to extend my warmest gratitude to my family, my brothers and sisters, Mpozi Nyaweza, and Nadine Bisimwa, who endured at least three years of apparent neglect with patience and courage. I also wish to thank Citoyens Kahegeshe, Mongane, Maheshe and their families for so much moral support. I also remember those institutions and students who willingly took part in my experiments.

Finally, I wish to thank Longman Publishers for allowing me to use a set of pictures from Byrne's (1967) Pupil's book in one of my experiments.

Edinburgh, September 1987

## Abstract

This study is a cross-sectional micro-analysis of the Interlanguage (IL) of 163 adult Zairean learners of English at three levels of proficiency. The analysis is based on elicited performance data. The informants are all speakers of French and at least two Bantu languages. By the time this study took place, the informants had had between three and six years of exposure to English, through formal instruction only. The data were collected during a four-month period of fieldwork including a pilot study and extensive observation of the informants by the experimenter.

Two main areas of linguistic variability and change in the interlanguage grammar are intensively scrutinized to describe the mapping of form-function relationship, within the framework of both contextual and free variability. These areas are: 1) the use of reference in second language acquisition and 2) the interpretation of infinitival complements under certain semantic constraints, or the influence of language universals. Additionally, the study provides an extensive analysis of some relevant features of the learner's developing grammars. The study also examines the effect of individual learner factors on rate and success in  $L_2$  learning.

From these analyses, a number of issues have been elucidated. Firstly, the study has shown that a cross-sectional study of a group of  $L_2$  learners of English, at different levels of proficiency, can be referred to in order to locate learners on an IL continuum. Secondly, movement along this continuum is systematic and variable at any point in time. Thirdly, the production of certain linguistic forms occurs before these forms are well understood and, or allocated to their appropriate functions. Thus, functional growth of these forms can be examined by looking at the learners' IL performance on different tasks. Fourthly, IL development obeys certain universal linguistic properties, to the extent that semantic and pragmatic constraints appear to influence the acquisition of certain IL subsystems, as in the case of complex sentences. Fifthly, in this multilingual setting, transfer appears to originate primarily from French (the first  $L_2$ ), rather than the learner's actual mother tongue. Sixthly, in a formal instructional environment, length of exposure to the  $L_2$  and the learner's degree of involvement in the learning task are more likely to lead to faster rate and success in SLA, than mere motivation or general favourable attitudes.

To conclude, the study shows the relevance of procedural approaches to learner language varieties for SLA theory, classroom-oriented research and educational policy.

## TABLE OF CONTENTS

<b>1 General Introduction and Background to the Study</b>	<b>1</b>
1.1 Theoretical Introduction	1
1.1.1 Sociolinguistic and Pedagogic Background of the Investigation	5
1.2 Structure of the Thesis	12
<b>2 A Review of Literature in SLA</b>	<b>14</b>
2.1 Morphemes and Other Syntactic Structures Studies	14
2.1.1 Orders of Morpheme Acquisition or Accuracy and the Notion of Language Learner's System	14
2.1.2 Methodological Problems in Morpheme Acquisition / Accuracy Studies	16
2.1.3 The Mother Tongue Influence Hypothesis	19
2.1.4 Auxiliary Studies and the Notion of Stages	20
2.1.5 More Advanced Syntactic Structures Studies	22
2.2 Theoretical Models for Second Language Acquisition	25
2.2.1 Krashen's Model	25
2.2.1.1 The Natural Order Hypothesis	25
2.2.1.2 The Monitor Hypothesis	27
2.2.1.3 The Acquisition / Learning Distinction Hypothesis	27
2.2.1.4 The Input Hypothesis	30
2.2.1.5 The Affective Filter Hypothesis	32
2.2.2 The Acculturation Model	33
2.2.3 The Nativization Model	36
2.2.4 Accommodation Theory	37
2.2.5 The Universal Hypothesis and SLA	40
2.2.5.1 Universal Grammar	41
2.2.5.2 Typological Universals	44
2.2.5.3 Universal Constraints on Interlanguages	46
2.2.6 Variability Analysis Models	53
2.2.6.1 Labov's Quantitative Paradigm	54
2.2.6.2 The Additive Model	55
2.2.6.3 The Multiplicative Model	57
2.2.6.4 The Logistic Model	60
2.2.6.5 The Variable Competence Model	61
2.2.7 Implicational Scale Model	63
2.3 Relevance of the Review of Literature to the Present Study	66
2.4 Summary and Conclusion	67
<b>3 Some Theoretical Considerations on the Aspects of IL under Investigation and Form-Function Approach</b>	<b>69</b>
3.1 Introduction	69
3.2 A Syntactic Description of the Article System	70
3.2.1 The Definite Article <i>the</i>	70
3.2.2 The Generic Article ( <i>the, a, 0</i> )	73
3.2.3 The Indefinite Article <i>a, an</i>	76

3.3	Some Theoretical Considerations on the Semantics of Articles	77
3.3.1	The Notion of Specific vs Non-specific Reference	79
3.3.2	The Notion of Shared Knowledge	81
3.3.3	Types of Specificity of Discourse Referents	82
3.3.3.1	Specificity by General Uniqueness	82
3.3.3.2	Specificity introduced Conversationally	83
3.3.3.3	Specificity by Entailment	84
3.3.4	A Few Problems in Definite and Indefinite Reference	85
3.3.4.1	The Ambiguity of Some Indefinite Expressions	85
3.3.4.2	A Problem with Entailment	86
3.3.4.3	Non-specific Reference to Particular Referents	87
3.3.4.4	Generic and Indefinite Expressions	88
3.3.4.5	Definite Expressions with Relative Clauses	88
3.3.5	Definite and Indefinite Reference in Second Language Acquisition	89
3.3.6	Studies of Reference in L <sub>1</sub> and SLA	93
3.4	Some Theoretical Considerations on the Infinitival Complementation	97
3.4.1	The Principle of Minimal Distance [PMD]	101
3.4.2	The Dative/Agent Principle	104
3.5	Some Theoretical Considerations on the Form-Function Approach	106
3.6	Summary	108
<b>4</b>	<b>General Experimental Design and Data Collection</b>	<b>109</b>
4.1	Testing Instruments and Directions for Fieldwork	109
4.2	Requirements of the Tasks	111
4.3	Task Types and Test Items	111
4.3.1	A Modified Cloze test for the Article System (74 items)	112
4.3.2	Sentence Interpretation for Infinitival Complements	114
4.3.3	Picture Composition Task	116
4.3.4	Questionnaire on Attitude and Motivation	116
4.4	Administration of the Tasks	118
4.4.1	Learner's Avoidance Strategy	118
4.4.1.1	Administrative Problems	119
4.4.2	Equipment	120
4.5	The Pilot Study	120
4.5.1	Aim	120
4.5.2	Population Sample	120
4.5.3	Administration of Testing Instruments	123
4.5.3.1	Modified Cloze Test	124
4.5.3.2	Sentence Interpretation	126
4.5.3.3	Written Composition Tasks <see Appendix IV>	129
4.5.3.4	Attitude and Motivation Survey	130
4.6	The Main Study	131
4.6.1	Assumptions	131
4.6.2	Subjects	132
4.6.3	Administration of the Tasks	132
4.6.3.1	At the Secondary Level	132
4.6.3.2	At the University Level	133
4.6.3.3	At the Teacher's College	134
4.7	Summary	135

## TABLE OF CONTENTS

<b>5 Results of the Use of Reference in SLA</b>	<b>137</b>
5.1 Group Trends	137
5.1.1 Group Range Variability	137
5.1.2 Analysis of Variance and Scheffe Tests	142
5.2 Individual Trends	151
5.2.1 Scalogram Analysis for Articles	151
5.2.2 Scalogram Analysis for Articles within 3 Linguistic Environments	161
5.3 Learner Language Behaviour	164
5.3.1 The Use of Definite and Indefinite Reference In Obligatory Contexts	165
5.3.2 The Use of the Definite and Indefinite Reference in Ungrammatical Contexts	165
5.3.3 The Use of Definite and Indefinite Reference in Three Linguistic Environments	171
5.4 Developmental Stages in the Use of Reference	174
5.4.1 The Notion of Developmental Stages in SLA	174
5.4.2 Determining Developmental Stages in the Use of Reference	174
5.5 Discussion	182
<b>6 Results of the Investigation of the Complement Sentences and Universal Topicality Hierarchies</b>	<b>186</b>
6.1 Group Trends	186
6.1.1 Group Range Variability	186
6.1.2 Analysis of Variance: ANOVA and Scheffe Tests	192
6.1.3 Scalogram Analysis	199
6.2 Individual Trends in the Performance on Complement Structures	205
6.2.1 Implicational Scales for Verb Types	205
6.2.2 Implicational Scales for Universal Topicality Hierarchies	212
6.3 Learner Language Behaviour	213
6.3.1 <i>Tell</i> Sentences	215
6.3.2 <i>Ask</i> Sentences	218
6.3.3 <i>Promise</i> Sentences	223
6.4 Discussion	226



TABLE OF CONTENTS

<b>7 Assessing Global L<sub>2</sub> Language Proficiency</b>	<b>233</b>
7.1 Introduction	233
7.1.1 A Brief Review of Literature	234
7.2 Some Theoretical Considerations	236
7.2.1 The Notion of Mean Utterance Length (Brown and Fraser, 1964)	236
7.2.2 The Notion of T-Unit (Hunt, 1965)	237
7.2.3 Bever's (1972) Speech Perception Strategy	239
7.2.4 The Notion of Information Blocks (Pepinsky, 1974)	241
7.2.5 The Notion of Communicational Capability Index (CCI)	242
7.3 Design	249
7.4 Subjects	252
7.5 Materials and Procedures	253
7.6 Performance Analysis: An Information Processing Perspective	254
7.6.1 Method of Analysis of Data	254
7.6.2 Error Identification	255
7.6.3 Classification of Errors: Top Ten List	256
7.6.4 Quantitative Analysis of Data	260
7.7 Learner Language Behaviour: A Re-appraisal	268
7.8 Learner Language Behaviour: Noun Phrase Related Features	272
7.8.1 Definite Article	272
7.8.2 Indefinite Article	275
7.8.3 Personal Pronouns	277
7.8.4 Lexical Choice	279
7.8.5 Relative Pronouns	281
7.8.6 Demonstratives	282
7.8.7 Quantifiers	283
7.9 Learner Language Behaviour: Verb Phrase Related Features	284
7.9.1 Transitive verbs	284
7.9.2 Intransitive Verbs	290
7.9.3 Copula <i>Be</i>	299
7.9.4 Auxiliaries	300
7.9.5 Subordination	302
7.9.6 Infinitival Complementation	304
7.10 Learner Language Behaviour: Prepositional Phrase Related Features	304
7.10.1 Learner's Performance on the Prepositional Phrase	306
7.10.1.1 Place	306
7.10.1.2 Time	310
7.10.1.3 Post-modifier of noun phrase <i>of</i>	312
7.10.1.4 Other Adjuncts: <i>for, with, by</i>	312
7.11 Discussion	314
7.12 Summary	325

## TABLE OF CONTENTS

<b>8 The Relationship Between The Learner Factors and Language Proficiency</b>	<b>327</b>
8.1 Introduction	327
8.2 A brief review of studies on the influence of the learner factors on Language Proficiency.	328
8.3 Experimental Design	331
8.4 Subjects	334
8.5 Materials and Procedure	334
8.6 Results	334
8.6.1 Correlations and Factor Analysis	335
8.6.2 Analysis of Variance and Scheffe Tests	343
8.7 Discussion	350
8.8 Summary	360
<b>9 Summary, Significance of the Study and Conclusion</b>	<b>362</b>
9.1 Introduction	362
9.2 Features of the Current Study	363
9.3 Significance of the Findings of the Present Study	364
9.3.1 Significance of the findings for L <sub>2</sub> Acquisition Research	365
9.3.1.1 Linguistic Criteria for Transferability	366
9.3.1.2 Psycholinguistic Criteria for Transferability	366
9.3.1.3 Socio-psychological Criteria for Transferability	367
9.3.2 Significance of the Findings for L <sub>2</sub> Pedagogy and Educational Policy	371
9.3.2.1 Significance of the Findings for L <sub>2</sub> Pedagogy	372
9.3.2.2 Significance of the Findings for Educational Policy	373
9.4 Future Directions for Research	377
9.4.1 Future Research in the Area of Reference in SLA	377
9.4.2 Future Research on the Role of Language Universals	377
9.4.3 Future Research on L <sub>2</sub> Development Index	378
9.4.4 Future Directions for Research on the Learner Factors	379
9.5 Conclusion	380

## BIBLIOGRAPHY

## Appendix I Results of the Performance Analysis: Chapters 5 and 6

### List of Tables in Appendix I

Table	Page
Table A Implicational scale for the Pupils' use of definite reference <i>the</i> in 3 environments	i
Table B Implicational scale for the undergraduates' use of definite reference <i>the</i> in three environments	ii
Table C Implicational scale for the T-Teachers' use of definite reference <i>the</i> in three environments	iii
Table D Implicational scale for the pupils' use of indefinite reference <i>a, an</i> in three environments	iv
Table E Implicational scale for the undergraduates' use of indefinite reference <i>a, an</i> in three environments	v
Table F Implicational scale for the T-teachers' use of indefinite reference <i>a, an</i> in three environments	vi
Table G Implicational scale for the pupils' use of definite reference <i>the</i> <sub>2</sub> in three environments	vii
Table H Implicational scale for the undergraduates' use of definite reference <i>the</i> <sub>2</sub> in three environments	viii
Table I Implicational scale for the T-teachers' use of definite reference <i>the</i> <sub>2</sub> in three environments	ix
Table J Developmental stages among 3 article types within 3 environments	x
Table K Scheffe tests: a 2-way interaction between verb types and the semantic features of the main clause Subject	xii
Table L Scheffe tests: a 2-way interaction between verb types and the semantic features of the main clause Object	xiii
Table M Scheffe tests: a 2-way interaction between Object and Subject	xiv
Table N Guttman scale for Subject and Object animacy features	xv
Table O Implicational scales for individuals: pupils	xvi
Table P Implicational scales for individuals: undergraduates	xvii
Table Q Implicational scales for individuals: postgraduates	xviii
Table R Results of <i>T'ell</i> sentences: NP <sub>1</sub> and NP <sub>2</sub> have equal status	xix
Table S Results of <i>T'ell</i> sentences: NP <sub>2</sub> is higher on the topicality hierarchy: convergence	xx
Table T Results of <i>T'ell</i> sentences: NP <sub>1</sub> is	

higher on the topicality hierarchy: conflict	xxi
Table U Results of <i>Ask</i> sentences: NP <sub>1</sub> and NP <sub>2</sub> have equal status	xxii
Table V Results of <i>Ask</i> sentences: NP <sub>2</sub> is higher on the topicality hierarchy: convergence	xxiii
Table W Results of <i>Ask</i> sentences: NP <sub>1</sub> is higher on the topicality hierarchy: conflict	xxiv
Table X Results of <i>Promise</i> sentences: NP <sub>1</sub> and NP <sub>2</sub> have equal status	xxv
Table Y Results of <i>Promise</i> sentences: NP <sub>1</sub> is higher on the topicality hierarchy: convergence	xxvi
Table Z Results of <i>Promise</i> sentences: NP <sub>2</sub> is higher on the topicality hierarchy: conflict	xxvii
Figure A Overall Patterns for <i>Tell</i> sentences	xxviii
Figure B Overall Patterns for <i>Ask</i> sentences	xxix
Figure C Overall Patterns for <i>Promise</i> sentences	xxx
<b>Appendix II Testing Instruments: Modified Cloze Test</b>	xxxii
<b>Appendix III Testing Instruments: Sentence Interpretation Task</b>	xxxiv
<b>Appendix IV Testing Instruments for Composition</b>	xxxviii
<b>Questionnaires for the Attitude-Motivation Survey</b>	xl

### **Abbreviations appearing in the text**

ACEW = Average clause embedding weight

ADLA = Authoritative disciplined learning assistant (our idea of the teacher in a learner-centred approach)

CCI = Communicational capability index

D = Definite reference

EFIB = Error-free Information block

ELT = English Language teaching

HK = Assumed hearer's knowledge

Interlanguage = IL

L<sub>1</sub> = First language

L<sub>2</sub> = Second language

MDP = Minimum Distance Principle

MUL = Mean utterance length

NP = Noun phrase

PP = Prepositional phrase

SE = Standard English

SLA = Second Language Acquisition

SR = Specific reference

TL = Target language

UG = Universal grammar (theory)

VP = Verb phrase

## List of figures in text

Figure no.		Page
3.1	A Semantic Configuration of the Article system	91
5.1	Percentage of subjects scoring > 80% correct for each article type	142
5.2	Mean scores for each level on four article types	143
6.1	Histogram for three verbs	192
6.2	<i>Tell</i> sentences equal status of NP <sub>1</sub> and NP <sub>2</sub>	216
6.3	<i>Tell</i> sentences convergence of NP <sub>1</sub> and NP <sub>2</sub>	217
6.4	<i>Tell</i> sentences conflict between NP <sub>1</sub> and NP <sub>2</sub>	217
6.5	<i>Ask</i> sentences - equal status	219
6.6	<i>Ask</i> sentences - convergence	220
6.7	<i>Ask</i> sentences - conflict	222
6.8	<i>Promise</i> sentences - equal status	224
6.9	<i>Promise</i> sentences - convergence	225
6.10	<i>Promise</i> sentences - conflict	226
7.1	Speech Perception Model (Bever, 1972)	240
7.2	An illustration of the Informative Display in deep structure	244
7.3	Percentage of errors per level for each major major sentence constituent	258

## List of Tables in Text

Table		Page
2.1	The Implicational Scale Model (adapted from Andersen 1978)	64
4.1	Elementary and Intermediate Group Means Differences	126
5.1	Group Range Variability of definite <i>the</i>	138
5.2	Group Range Variability of definite article <i>the</i> <sub>2</sub>	139
5.3	Group Range Variability of article <i>0</i>	140
5.4	Group Range Variability of the indefinite articles <i>a, an</i>	141
5.5	Overall Performance per article type: All levels combined	144
5.6	ANOVA results for four articles in three environments by level	147
5.7	Scheffe tests for Comparison between articles	148
5.8	Scheffe tests for Comparison between pairs of environments	149
5.9	Scheffe tests for articles within environments	150
5.10	Guttman Scale for 3 articles	153
5.11	Proportion of Subjects reaching the 80% criterion	154
5.12	Implicational Scale for 2nd year teacher trainees	156
5.13	Implicational Scale for the Pupils' use of reference	158
5.14	Implicational Scale for the Undergrads use of reference	159
5.15	Implicational Scale for the T-Teachers use of reference	160
5.16	Most favoured Orders of Accuracy for Article Types	161
5.17	Implicational Patterns for use of definite Reference <i>the</i>	162
5.18	Implicational Patterns for the use of indefinite reference	162
5.19	Implicational Patterns for the use of definite reference <i>the</i> <sub>2</sub> at three levels	163
5.20	Percent Correct use of the articles in obligatory contexts	164
5.21	Percent Correct Realization of the article system in both obligatory and Ungrammatical contexts	166
5.22	Correct and Incorrect uses of <i>the</i> in three linguistic environments	169
5.23	Correct and incorrect uses of <i>the</i> <sub>2</sub> in three linguistic environments	170
5.24	Correct and incorrect uses of <i>a, an</i> in three linguistic contexts	171
5.25	Correct use of each article within three linguistic environments by levels	173

5.26	Developmental stages among three articles within three environments: bimodal implicational patterns	177
6.1	Group Range Variability of <i>Promise</i> Verbs	187
6.2	Group Range Variability of <i>Ask</i> Verbs	188
6.3	Group Range Variability of <i>Tell</i> Verbs	189
6.4	ANOVA results for 3 verbs, 3 subjects and 3 objects types by levels	195
6.5	Scheffe tests for comparison between 3 verbs	196
6.6	Scheffe tests for comparison between verbs and levels	197
6.7	Scheffe tests for 3-way interaction VSO	200
6.8	Guttman Scale for 3 verb types	202
6.9	3 left-most combinations of features for NP <sub>1</sub> and NP <sub>2</sub>	204
6.10	3 right-most combinations of features for NP <sub>1</sub> and NP <sub>2</sub>	204
6.11	Combination of features in increasing order of difficulty	205
6.12	Implicational scale for the pupils' performance on three verbs	207
6.13	Implicational scale for the undergrads' performance on three verbs	209
6.14	Implicational scale for the teachers' performance on three verb types	210
7.1	Text sample used for error analysis	240
7.2	Frequent Error types: Top ten list	256
7.3	Frequency of Errors per major sentence constituent and levels	257
7.4	Pearson Correlation between 4 performance variables	261
7.5	ANOVA results for average no. of clusters per information block by levels	262
7.6	Scheffe tests for average no. of clauses per information block with levels	262
7.7	ANOVA results for average clause embedding weight by levels	263
7.8	Scheffe tests for average clause embedding weight with levels	264
7.9	ANOVA results for error-free information blocks by levels	265
7.10	ANOVA results for error-free information blocks by levels with average clause embedding weight as covariant	266
7.11	Noun phrase performance features	273
7.12	Verb Phrase performance features	284
7.13	Verb phrase performance features: intransitive verbs	291
7.14	Overall verb phrase performance features: all features	300
7.15	Prepositional phrase performance features	306
8.1	Pearson Correlation matrix for attitude/motivation survey and two language tasks	336
8.2	Unrotated principal factor solution (PA2) with	



	iterations	337
8.3	Factor matrix (Varimax) rotation with iterations	338
8.4	Pearson correlation matrix for attitude/motivation survey and four language tasks	340
8.5	Unrotated principal factor solution (PA2) with iterations: four subtests and four L <sub>2</sub> performance tasks	341
8.6	Factor matrix (Varimax rotation with iterations): four subtests and and four performance tasks	341
8.7	ANOVA -Instrumental motivation by level of training in English	344
8.8	ANOVA -Sentimental motivation by level of training in English	345
8.9	ANOVA -Attitudes towards English language/culture by levels of training in English	346
8.10	ANOVA -Desire to learn English by level of training in English	347
8.11	ANOVA -Performance on the modified cloze test by level of training in English	348
8.12	ANOVA- Performance in the sentence interpretation task by levels	348
8.13	ANOVA -Average clause embedding weight by level	349
8.14	ANOVA -Proportion of Error-free Information blocks by level	350

## CHAPTER 1

### GENERAL INTRODUCTION AND BACKGROUND TO THE STUDY

#### 1.1. Theoretical Introduction

The speculation that a natural syllabus exists for first and second language acquisition and possibly for language acquisition in general, is one of the most insightful proposals in current linguistic research, in terms of its relevance to second language acquisition theory and second language pedagogy. Thus, Corder (1967) suggests that learners have a 'built-in syllabus' for the learning of any particular second language. Since then, the hypothesis has been extensively investigated by many other researchers who seem to agree with the idea that the natural cognitive processes of learning, when faced with a particular body of data, determine the sequence of creating that cognitive structure which is called the grammar of a language (Dulay and Burt, 1973, 1974; Bailey, Madden and Krashen, 1974; Larsen-Freeman, 1975; Hatch and Wagner-Gough, 1976; Hyltenstam, 1977; Corder, 1973, 1978, 1981).

In the early sixties the role of the mother tongue in second language acquisition had come to be established as axiomatic. As a result, Contrastive Analysis (Lado, 1964) was considered a most successful way of applying linguistic theory for classroom teaching purposes. Within less than a decade, Contrastive Analysis and the hypothesis of mother tongue interference had been attacked from various standpoints. Despite the severe criticism both Contrastive Analysis and mother tongue interference hypothesis have suffered, they have never been completely floored (Wilkins, 1968; Newmark and Reibel, 1970; and also Hakuta, 1974; Wode, 1976)

The concept of Interlanguage (Selinker, 1972) and the related notions of

Approximative Systems (Nemser, 1971) and Transitional Competence (Corder, 1967) have contributed to a better understanding of what goes on during a language learner's career. Such notions have indeed further diminished without, however, annihilating the mother tongue interference hypothesis as a determining factor of second language acquisition. Thus, it can be observed that all the strictures against the mother tongue interference hypothesis have not been strong enough to uproot its foundations.

On the other hand, the emergence of abundant research in intra-lingual or developmental errors in language acquisition has provided grounds for reasonable argument that a natural syllabus exists for first and second language, and that all language learners follow roughly the same route of development (Brown, 1973; De Villiers and De Villiers, 1973; Burt and Dulay, 1973, 1974a, b; Ervin-Tripp, 1974; Bailey, Madden and Krashen, 1974; Duskova, 1969; Richards, 1970; Ellis 1984, 1986)

Such findings have also indicated that, while there is a good deal of argument about the degree of systematicity, the move from the beginning stages of language learning to later ones follows a sequence which is not random (Hatch, 1978). Similarly, it had been demonstrated that the learner's age does not interfere significantly with the 'natural' route of acquisition (Krashen, 1973; Krashen, Sferlazza, Feldman, and Fathman, 1976), a finding which has invalidated earlier claims that the completion of cortical lateralisation at puberty was the cause for adults' difficulties in acquiring second language (Lenneberg, 1967). It is generally understood, however, that the learner's age does involve a difference in learning style and in the rate of L<sub>2</sub> acquisition (Richards, 1978). One of the major characteristics of the language of the language learner is that it is highly variable for both children and adults. In

other words, while the learners' interlanguage may show systematic changes, this system is not an invariant one for all learners. Even within one particular learner's data, there appears to be a great deal of overlap as he moves from one stage to another along the developmental continuum. In spite of the variability observed in the learners' approximative systems there are also striking similarities which allow us to talk of the interlanguage as evidence that basic processes are at work in the acquisition of a second language.

In Selinker's (1969, 1972) sense, Interlanguage refers to a language system which he believes is intermediate between the learner's mother tongue and the target language on the grounds that it shows some formal characteristics of both. Early work in Interlanguage was done therefore along the lines of the Language Transfer phenomenon. This involved the idea that Interlanguages were unstable and seen as continua in which items from one system are gradually replaced by items from another, that is a restructuring continuum. Selinker also points out that many  $L_2$  learners fail to reach target language competence or the end of the IL continuum. They stop learning when their interlanguage contains some rules which are at variance with those of the target language system. Selinker refers to this as *fossilization* and he speculates that this phenomenon cannot be remedied by further instruction. However, fossilized structures may not be persistent. Sometimes the learner may succeed in producing the correct form, but when he is focused on the meaning (e.g. in casual style, or when dealing with a challenging topic) he will *backslide* towards his true IL norm. According to Selinker and Lamendella (1978), fossilization may be due to both internal and external factors. It can occur both because the learner believes that he does not need to develop further his IL any further to meet his communicational needs; or because the neural structure of his brain has reached a critical period; and thus, his hypothesis-testing mechanisms fail

to operate satisfactorily.

The view of IL as a purely restructuring continuum has been questioned by some researchers (Burt and Dulay, 1974) who have provided some evidence suggesting that the interlanguages of second language learners do not necessarily show transfer from the mother tongue. Further criticism has come from Corder (1978) who observes that in such cases, movement along the continuum implies that the overall complexity of the language remains the same at any point along the continuum. Corder, then proposes that learning involves the increasing complexification of one linguistic system as it moves towards another. This Corder refers to as a *recreation* or *developmental* continuum. The concept of hypothesis testing has been used to explain how the L<sub>2</sub> learner progresses along the developmental continuum, in much the same way as it was used in L<sub>1</sub> acquisition studies. Corder saw the making of errors as a strategy or evidence of learner-internal processing. Foreign language learning is admittedly a combination of both types of continua (i.e. restructuring and developmental continua).

Recent approaches to the language of the L<sub>2</sub> learner have led to the recognition of a new type of English, viz. 'Learner English' as a variety of English (i.e. from a functional or socio-cultural perspective). One of the characteristics of Learner English is that it contains at least some rules that differ considerably from the language spoken as a native language by the majority of the population of Great Britain, Canada, the United States and Australia. Another important aspect of Learner English is that it is often primarily learnt inside the classroom, rather than in more normal communicative situations. Learner English is a variety of English in its own right. However, it is by no means a homogeneous variety. It differs according

to:

1. The sociolinguistic and socio-cultural background of the speakers (e.g. French, Chinese, Swahili, Yoruba, Danish, Spanish, etc.)
2. The nature of the learning situation (e.g. classroom teaching only, or classroom teaching in addition to informal exposure); and,
3. The language competence level of the learner

In this respect, Learner English involves typically the variety used between an L<sub>2</sub> speaker (e.g. Swahili-English, Italian-English) and a native speaker of English (or a more competent L<sub>2</sub> speaker of English such as an L<sub>2</sub> teacher). It also includes the learner varieties used as a means of communication among speakers from different speech communities, viz. the varieties often referred to as 'New Englishes' (Platt, Weber and Ho, 1984).

The recognition of the linguistic systems of second language learners as a well defined area of research and theory within the mainstream of language study has come about as a result of considerable research activity over the last two decades or so. Several models of second language acquisition have been proposed (Ellis, 1986). In the next chapter a review of the literature in SLA is presented.

### **1.1.1. Sociolinguistic and Pedagogic Background of the Investigation**

When Zairean children start their formal education, they bring with them (among other things) the experience or background of two or three Bantu languages to their task of learning French, and English at a later stage. In their homes and other areas of their social life, Zaireans use either their tribal language (i.e. their real mother tongue), or one of four lingua francae in common use in the country viz. Kikongo, Lingala, Swahili and Tshiluba. Thus, in Kinshasa city, for instance, every Zairean adult is assumed to have a good

command of at least three Zairean languages. Children born in Kinshasa, therefore, undergo considerable pressure from their parents to learn either their tribal language or at least the lingua franca spoken in their region of origin. This appears to be a difficult (and relatively unsuccessful) task since in most cases there is no significant literacy in any of these languages. On the other hand, children are aware that they can fulfil most of their vital communication needs through the medium of Lingala, i.e. the lingua franca which is dominant in Kinshasa. Some children do indeed learn their tribal language as far as they can, if only in order to achieve sufficient receptive competence. The latter is important to the extent that it is required during certain important social gatherings such as funerals, marriages and other traditional festivities involving the attendance of all the members of the tribe.

But children whose parents were born in Kinshasa themselves find it almost impossible to learn their original tribal language (through lack of exposure or motivation) and restrict themselves to the knowledge of Lingala. The relevant point here is the fact that for neither group of children is French the language used at home. At school, all Zairean children are confronted with the difficulty of studying French as a subject and using it as a medium of instruction. A very good command of French is important because it can affect academic performance, and, therefore allow access to higher education and upwards social mobility.

However, overall communicative competence in French remains very modest among Zairean children, owing to poor teaching materials and methodology, as well as restricted input and functional use. On the other hand, French is the official language, that is, the language of administration, banking, international trade and business correspondence (but with an increasing

possibility of encroachment of this function by English). French has, therefore, a good deal of prestige even though it does not intervene much in the Zairean's private life.

English itself was introduced in the curriculum only in the late fifties, along with the implementation of the 1958 'Metropolitan' Programme. At first, English was taught as a subject only at secondary school, from the third year up till the sixth (final) year. Then, in the early seventies it was introduced in most faculties at the undergraduate university level. It is not clear, however, what the Belgian colonial administration's aim was in introducing English in the Zairean curriculum (rather than Dutch, for instance). As a matter of fact, the Belgian educational policy purported to provide only the strict minimum numbers of staff required to carry out low level duties within the Colonial administration network. More specifically, the Belgians were not interested in the development of a full-fledged educational system whereby the young people would be encouraged to move from the primary, through secondary, to the tertiary level (where their knowledge of English is most useful). In their view, such a system would have had the effect of creating an elite of the type that emerged in former French colonies, a fact which in some cases led to cultural assimilation and anomie.

Instead, all that the colonial administration seemed to be concerned with was the provision of subliterate staff, such as clerks, catechists, instructors in church schools, foremen in the plantations etc. In fact, over decades of colonial rule (roughly between 1890s and 1940s) there had been a protracted controversy among successive Belgian governments as to which language, either French or Dutch, would be retained as the official language for Zaire (then known as the Congo). Partly, this controversy was due to the continual



linguistic quarrel between Belgian French speaking Walloon and Dutch speaking Flemish groups. In 1936, French was eventually agreed upon as the Colony's official language, presumably in view of its wider potential for international communication. With hindsight, it can be suggested that English was introduced too with similar assumptions in perspective. Consequently, after Zaire's independence in 1960, English was extended to all of the secondary schools, and at present it is also taught at the university undergraduate level, as well as in certain private evening schools.

In recent years, there has been a keen interest in learning English measured by an increase in favourable attitudes among Zairean students (this will be shown later in section 8.6, in relation to the results of our own survey). The most obvious reason for this increase is the fact that a similar trend in the demand for 'bilingual' staff is known to have arisen in many multi-national companies as well as other institutions involved with the Zairean economy and trade. There is also the fact that at the university, Zairean students have found themselves involved with the struggle of operating in English, using specialist subject materials laid out in unfamiliar modes. As a result, a number of teacher training colleges have been created to supplement work hitherto done in universities' English departments. Both institutions have indeed been under constant pressure to provide teachers of English (and other languages) in great numbers; and, they have been turning scores of new teachers of English every year. The quality of teaching/learning English, however, remains substantially modest. One reason for this is the lack of clear guidelines in the curriculum in relation to the needs and objectives of the teaching of English. During my fieldwork I became aware that, in fact, there were no specific guidelines indicating what the content of the English course should consist of. High ranking officials in the Ministry of Primary and Secondary Education

acknowledged the fact that 'apart from the choice of textbooks, which has so far been left to the whims of technicians of the Belgian Cooperation, every teacher is free to teach his pupils what he likes, the way he likes' (Makundi, B., an adviser to the State Commissioner, or Minister of Education: personal communication).

Another reason is that the teacher's morale has fallen to a very low level as a result of the teacher's deteriorating living conditions. This has led to the teachers' widespread disenchantment or disillusionment, a phenomenon which has been referred to by educational psychologists as the 'Burn-Out'. The lack of clear guidelines in the curriculum is even more severe at the university level where the textbooks themselves are non-existent. At this level, every teacher is expected to devise his own syllabus. This often amounts to nothing else than compiling a number of reading passages taken, at random, from outdated manuals which are putatively related to the students' specialist subject. The teachers' difficulties are further compounded by a number of administrative as well as pedagogic problems, including overcrowding in the classroom, and the lack of resource centres etc. The latter are of considerable importance because they could provide the teacher (both the newly trained and the experienced ones) with up-to-date information regarding new ideas on syllabus design principles and classroom management. In this respect, the situation of English teaching in Zaire can be described as chaotic and reflecting a general condition of mediocrity. If the situation were to change in the positive direction, then, the government ought to take a number of steps to the effect of committing more funds to curriculum development, the creation of resources centres, and, of course, the improvement of the teachers' living standards. More importantly, the government should be more inclined to base their decision on informed assessment of the educational system by specialists in the field, viz.

educationalists, applied linguists and sociolinguists.

Of course, all this does not mean that learning English cannot or has not taken place at all in spite of these adverse conditions. It would be totally wrong to argue along those lines. The suggestion worth making is that it is possible to improve upon the state of affairs referred to above. This would involve more teachers being well trained and a variety of materials being placed at their disposal. If this was the case, it is hoped that a substantial increment would be induced in relation to the amount of knowledge being acquired by the Zairean students as well as the achievement of a higher degree of proficiency at a faster rate.

Indeed, there have been attempts to improve the English Language teaching (ELT) situation in Zaire. But these have remained sporadic, unco-ordinated at a higher level or official level, and potentially, misguided.

Teachers of English in Zaire too are known to have made some attempts to create materials of their own to suit the needs of their students. One case in point is the ongoing projects undertaken by a group of EFL teachers, in order to provide the secondary school system with a comprehensive set of six textbooks to be used from the first year to the sixth of secondary school (Kasama, Ndoma, Ntahwakuderwa, and Swekila, forthcoming). The project was initiated following the recommendations of the 'New Programme' for Zairean secondary schools (cfr. letter DEPS/BCE001/0290/80 dated 07 August 1980). According to the New Programme 'the English course should be taught in such a way as to make pupils familiar with the sounds and some basic structures of the language, with a major focus on listening and speaking' (Rapport Final de Synthese, Nouveau Programme 07/08/1980 p.3).

Basically, the authors have adopted an approach which emphasises fluency, rather than accuracy, in two specific areas, viz. listening comprehension and speaking. One of the authors' main concerns has been to find a compromise solution between the objectives of the course as set out in the (rather structuralist) programme and the methodological principles suitable for a communicative approach-based presentation of the materials.

Assuming that good teaching materials could be eventually introduced, one should not jump to the conclusion that the learning of English will improve automatically for all learners. The reason is a simple one: that is, there is no direct, causal relationship between what is taught and what is eventually internalized as intake, by the learners. When exposed to language data, the latter are known to process and analyse it according to their own individually perceived psycholinguistic units, i.e. with reference to an hypothesized 'built-in' syllabus (Corder, 1967) of some kind. However the teachers' role is to facilitate the learner's task, and therefore they may be entitled to claim credit for much of what the learner eventually acquires. It is indeed true to say that the majority of learners are more likely to benefit from good teaching materials than ill-conceived ones or none. But, if anything, good teaching materials or techniques impel, not compel, learning or acquisition.

As far as both teachers and researchers are concerned, it remains a continuing task to be able to evaluate the learner's progress and success, although criteria for the latter are not always so easy to establish. Thus, a careful investigation of what the learner is doing (or, in spite of) the materials is important because it can give us an indication as to what has been learned and how. It has also been suggested that second language acquisition research will just tell us how *not* to teach whatever we want to teach (Lightbown, 1985).

As far as the present investigation, the question one has to answer at this stage is: 'What is it then, that this work intends to concern itself with?'. In the next section we shall provide an overview of the structure of the thesis.

## **1.2. Structure of the Thesis**

The purpose of this study is to examine form-function relationships in the interlanguage of Zairean French speakers learning English in the classroom environment. Because second language acquisition is best conceived of as a multi-dimensional process (Meisel, Clahsen and Pienemann, 1981), a number of rather different issues will be dealt with in the various chapters that constitute the thesis. The thesis includes nine chapters, a bibliography and appendices. Chapter One consists of a general introduction and presents the sociolinguistic and pedagogic backgrounds to the study. In its theoretical introduction, Chapter One provides a new outlook on the language of the language learner as a variety of English in its own right, encompassing both interlanguage and so-called New Englishes (Platt et al. 1984) within the same framework of  $L_2$  learners' variable competence. Chapter Two presents a review of the relevant literature in SLA, including earlier work in connection with the Morpheme Studies, the methodological problems caused by certain elicitation techniques as well as a brief account of certain models for SLA. Chapter Three presents some theoretical considerations on the aspects of interlanguage under investigation. These are: 1) the use of definite and indefinite reference in SLA; and 2) infinitival complementation, especially the relevance of such concepts as the Minimum Distance Principle (MDP) (Rosenbaum 1967; Chomsky 1969, 1972). Additionally, Chapter Three describes the form-function approach to learner language data, as a way of resolving some of the inadequacies inherent in the earlier form-form approach. Chapter Four provides the general methodological framework and the methods used for collecting data. Chapter Five presents the

results of the study of the use of reference in SLA. Both quantitative and qualitative analyses are conducted in order to map form to function relationships in the L<sub>2</sub> development. The findings reveal that roughly five developmental stages are evidenced in the learners' performance. Meanwhile such notions as 'specificity' and 'shared knowledge' appear to play a certain role in the IL development of the learners under investigation.

Chapter Six examines how both the syntactic and semantic/pragmatic factors affect the learner's interpretation of complex sentences, i.e. the *tell*, *ask*, *promise* sentence types. The focus in this chapter is on the discovery of regularities in the way languages (i.e. both natural languages and ILs) vary, and on the universal constraints and principles that underlie this variation.

Chapter Seven is concerned with the assessment of global L<sub>2</sub> proficiency. An L<sub>2</sub> communicational capability index is proposed as an improvement on the notion of T-Unit (Hunt, 1965), as a clause analytical technique. The proposed index consists of a reliable and readily applicable instrument which will enable every researcher or classroom teacher to monitor his students' L<sub>2</sub> development in a meaningful way. Additionally, this chapter presents a full account of learner-language behaviour, by analysing both 'errors' and 'non-errors' in writing data.

Chapter Eight addresses itself to the effect that such learner factors as motivation, attitude, and length of L<sub>2</sub> exposure can have on IL development and eventual success in the target language. Chapter Nine sketches a general summary of the main findings of the study. It points out their significance for SLA theory and practice. It also highlights certain areas of interest for further SLA research; and, finally, it provides some concluding remarks. The thesis also contains an extensive bibliographical section and appendices.

**CHAPTER 2****A REVIEW OF LITERATURE IN SLA**

For the purpose of the present study, a review of only a few studies or models (which appear to be most relevant) will be undertaken. The models included in the review are 1) the Acculturation Model, 2) the Nativization Model, 3) the Accommodation theory, 4) the Monitor theory, 5) the Universal hypothesis, 6) Variability Analysis Models, and 7) the Implicational Scale Model. What these models have roughly in common is the fact that they emphasise that second language acquisition is not restricted to settings where teaching is a necessary part of the input. They also attempt to show how and why different types of learning and different contexts of language use lead to particular results. More importantly, these characteristics of the learning environment may affect the rate but not necessarily the 'natural' route of L<sub>2</sub> development. First, I will review a few studies in connection with the morpheme studies. Secondly, I will present a brief account of some theoretical models for SLA. Thirdly, I will point out the relevance of the review of literature, and then I will provide a summary for Chapter Two.

**2.1. Morphemes and Other Syntactic Structures Studies****2.1.1. Orders of Morpheme Acquisition or Accuracy and the Notion of Language Learner's System**

Brown's (1973) work on child language acquisition is considered a major landmark in both first and second language research. In his well-documented longitudinal study of three children (Adam, Eve and Sarah) learning English as a first language, Brown focused on the acquisition of 14 grammatical English

morphemes. He scored, namely, the present progressive; the third person singular both regular and irregular; past tenses both regular *ed* and irregular; the copula and auxiliary *be*; the article *a* and *the*; the prepositions *in* and *on*; the regular and irregular plural; and possessive inflections.

The results revealed that the 14 morphemes were acquired in a similar order by these children. Brown's findings were subsequently confirmed in a cross-sectional study by De Villiers and De Villiers (1973). Thus, Brown (1973: 105-106) suggested that:

. . . children work out rules for the speech they hear, passing from levels of lesser to greater complexity, simply because the human species is programmed at a certain period in its life to operate in this fashion on linguistic input.

Following Brown, other workers wondered if the the same order of acquisition of these morphemes might be found if they looked at second language data. Burt and Dulay (1973, 1974) carried out a series of studies using the De Villiers model, i.e. accuracy rather than acquisition data. They developed a technique called the Bilingual Syntax Measure (BSM) to elicit spontaneous data from Spanish speaking children (1973), and later, from Chinese speaking ones (1974) all learning English as second language.

Burt and Dulay found an invariant order of morpheme accuracy for the two groups. This enabled them to formulate what is now known as the Creative Construction Process hypothesis (Burt and Dulay, 1974).

The Creative Construction Hypothesis assumes that the learning process is automatic, and each learner creates his language anew via innate language learning ability. The learner's Language Acquisition Device (LAD) reacts to



unrestricted input data to create the system for the learner. The invariant order of morpheme acquisition or accuracy is thought to be a reflection of that internal rule-making LAD.

### 2.1.2. Methodological Problems in Morpheme Acquisition / Accuracy Studies

Many studies were conducted following Brown's (1973) longitudinal data and using his 90% criterion scoring method. But as observed in Hakuta (1974), many obtained substantially different results. Hakuta collected speech samples from Uguisu, a 5-year-old Japanese speaker learning English as a second language. Hakuta scored roughly the same morphemes as in Brown's study. However, the ranks obtained by Brown, De Villiers and De Villiers were different.

On the other hand, Burt and Dulay's (1974) methodology (the BSM technique) was used in several studies; and, although some reached results similar to those found by Burt and Dulay, others led to somewhat contradictory claims about the hypothesised 'invariant order' of acquisition /accuracy.

One of the major criticisms leveled against the invariant order hypothesis concerned the scoring method. It was argued that the '90% correct in obligatory instances' criterion loses much information. For example, if we count all instances of *be* as copula, we may find that the learner has indeed produced the copula correctly in 90% of the obligatory cases. But by looking more closely at the data, we may realise that only copula form *is* appears in the data, as in, e.g. *this is, that is, it's*. There may be no examples of *am, are, was, or were*. Yet, according to the 90% criterion the learner is assumed to have acquired the copula.

Also, it is possible to score a morpheme as acquired whereas, in fact, its function has not been acquired yet. For example, we may find that the present

progressive *-ing* forms appear in 90% of the obligatory instances; but the learner may also use it somewhere else. Gough (1975) reports, for instance, that her Iranian child used *-ing* forms almost everywhere, even for imperatives! Thus, Mikes (1967) has suggested that we should not claim acquisition of various forms until those forms appear in contrast to others in their class. For instance, *-ing* would only be considered acquired if it was used appropriately in contrast to present past and future time etc. In the case of learners of French, article and adjective agreement rules would not be acquired until the learners contrasted masculine and feminine endings appropriately.

Finally, the 90% criterion failed to account for the observed variability within the group. It could not tell what acquisition /accuracy order was like below the criterion level, that is in the 0-90% range.

In view of the criticism raised in connection with the 90% criterion scoring method, Andersen (1977) proposes a scoring method which helps detect variability within the data, namely the Group Range Method. This method shows the percentage of subjects who used any particular morpheme 90-100% of the time, 80-100%, 70-100%, etc.

For example, Percentage of Subjects using 8 Verb Phrase Morphemes at 90%, 80% and 70% criterion (Andersen, 1977).

	COP	AUX	ING	PAST IRREG
Criterion	N=83	N=83	N=83	N=72
90-100 %	94	53	48	31
80-100 %	98	66	61	56

The group range method can also show the percentage of subjects who always

used the morphemes correctly, those who never used it correctly, and those who are in the middle range.

Andersen also suggests that the choice of morphemes was not done satisfactorily. In his view there is a danger in choosing them on the basis of findings from child language studies. For instance, if articles (definite, indefinite and 0 article) are lumped together as a single morpheme, it is not possible to observe the amount of variability which he found in his learners' performance on definite / indefinite, and 0 article. Andersen proposes that more information would be gained by analysing the three forms separately.

Most morpheme studies reveal that morphemes acquired early appear together and those acquired last cluster together. Thus, it would seem inaccurate to assume that once rank-ordered, the difference between the morphemes is the same; that is to say, the difference between, for example, morpheme 1 and 2 is not necessarily the same as that between 5 and 6 in the rank. Morphemes 5 and 6 may sometimes be better analysed if they are regarded as located at the boundary of two different groups, rather than just two subsequent morphemes in the rank-order (Rosansky, 1976).

Rosansky (1976) has also criticised morpheme accuracy orders found by Burt and Dulay (1973, 1974). She points out that the (1973) Sacramento sample obtained an order which differs substantially from the (1974) Chinese sample, implying that the findings reported here may well be an artefact of Bilingual Syntax Measure (BSM) elicitation technique. But Rosansky next discovers that the order of morpheme acquisition found by Cazden et al (1975), using observational rather than elicited procedures, correlated with Burt and Dulay (1973, 1974), Bailey Madden and Krashen (1974), and Larsen- Freeman (1976) etc. Rosansky then concludes that caution, still, should be exercised when

using rank-orders because the latter tend to blur variability among learners, and may affect the significance of correlations.

### 2.1.3. The Mother Tongue Influence Hypothesis

In his work in first language acquisition, Brown used the average length of a child's utterance in morphemes, the mean length of utterance (MLU) as a measure of language development. Using MLU Brown described language acquisition in terms of a series of stages. Other workers in child language studies followed this model, and then analysed their data by writing formal phrase structure (PS) and transformational (TG) rules at each of these stages (Brown, Cazden and Bellugi, 1968; Ravem, 1974; Bellugi, 1967).

Huang (1970) describes the acquisition of negatives and questions for Taiwanese child learning English. Originally, Huang hoped to use the MLU method but he soon found that this procedure could not be applied to his older child because the latter could produce very long utterances from the start. For example, *It is time to eat and drink*. Huang's solution was to divide his data into months of exposure to English and then construct formal rules within these stages.

Ravem (1974) compared the rules involved for the acquisition of negatives both in English and Norwegian. Ravem predicted that Norwegian learners of English would produce correct negatives in view of the similarities which exist in negation in these two languages. The only difference, however, is that in Norwegian, in the case of main verb negation, the negative particle is placed after the verb, e.g. *He like not the house* or *He comes not today*. Ravem was surprised to find that his learners, instead, produced negatives similar to those of the first language learners of English, e.g. *He not like the house* and *He don't like it*.

For wh-questions, Ravem's children Reidun and Rune used inversion at an early stage. They produced forms which could not be generated by Norwegian, but similar to sentences used by Brown's children Adam, Eve and Sarah. For instance, while Norwegian would predict forms like *Where live Tom?*, they produced *Where Tom live?*. But for copula yes / no questions, they used intonation as a question marker only occasionally (e.g. *You like ice cream?*) or transfer of the Norwegian rule (e.g. *Like you ice cream?*). Reidun consistently chose intonation while Rune chose the Norwegian inversion pattern.

Therefore Ravem concluded that neither the  $L_2 = L_1$  learning hypothesis (Huang, 1970; Burt and Dulay, 1974; Ervin Tripp, 1974) nor the contrastive analysis hypothesis of interference / transfer (Lado, 1957; Hakuta, 1974), in their extreme versions, were corroborated by his data.

#### 2.1.4. Auxiliary Studies and the Notion of Stages

Adams (1974) reports on a study based on the acquisition of the English auxiliary system in questions and negatives by Spanish speaking children. She suggests that her subjects' development could be described in terms of stages. She notes, however, that these stages are not discrete but, rather, show a good deal of overlap among them. She argues, then, that this notion of Developmental Stages could be extended to all learners.

Adams found the following stages in the acquisition of negation. First, the negative element appeared within the sentence - *no* directly before the main verb, as in e.g. *I no sing it*. *Not* occurred before predicates where a copula would be required in adult speech. In this early stage *don't* appeared only in a few 'pre-fabricated chunks' (Hakuta, 1974) as in e.g. *I don't know*.

In the second stage an increase in the use of *don't* plus main verb pattern

became prominent. *Don't* was overgeneralised and used as a negative marker in place of other auxiliaries (*doesn't, won't, can't*). Then *don't* became a more restricted marker and the modals appeared. *Can't* and *won't* were first to appear. But many learners still had not produced sentences with the AUX element *have + en*.

Cazden, Cancino, Rosansky and Schumann (1975) also found that their data of 6 Spanish speaking subjects could be analysed according to different stages for English negatives and questions. For instance, *wh*-questions were described in two stages. First the learners did not distinguish between simple and embedded *wh*-questions. In some cases both simple and embedded *wh*-questions were uninverted. In others, simple *wh*-questions were sometimes inverted, sometimes not. Then there was an increasing degree of inversion in *wh*-questions, with inversion extended to embedded questions. In the second stage, learners were able to differentiate between simple and embedded *wh*-questions.

Fathman (1975) analysed data to discover the relationship between certain aspects of the second language acquisition process and age. Fathman developed an oral elicitation instrument known as the SLOPE to assess the ability of non-native English speaking children to produce standard English morphology and syntax. The 200 subjects ranged between ages 6-15 learning English as a second language in American public schools.

The test consisted of twenty sub-tests: (1) affirmative-declarative, (2) articles, (3) present participle, (4) possessive, (5) present tense 3rd person singular, (6) comparative, (7) superlative, (8) present tense 3rd person irregular, (9) preposition, (10) past participle regular, (11) negative, (12) past participle irregular, (13) subject pronouns, (14) object pronouns, (15) possessive pronouns,

(16) plural irregular, (17) imperatives, (18) yes/no questions, (19) wh-questions, and (20) plural irregular.

Fathman was interested in establishing the relationships between age and (i) the rate of acquisition of the structures and (ii) the sequence of acquisition of the same grammatical forms.

The results showed that among children with the same amount of exposure older children scored higher (i.e. acquired faster) on the morphology and syntax sub-tests, whereas the younger children scored higher on phonology. However, no significant differences were found in the sequence or order in which these structures were produced by the learners. Thus, Fathman suggested that there is a difference in the rate of learning of English morphology, syntax and phonology in connection with age differences; but the developmental sequence basically remains the same at different ages.

Krashen, Sferlazza, Feldman and Fathman (1976) used the SLOPE to test the validity of Lenneberg's (1967) Critical Period Hypothesis. According to Lenneberg, natural and complete acquisition of language can only take place before puberty and that the processes of language acquisition in children and adults are quite different. Krashen et al found that the difficulty order was not significantly different between both children and adults learning English as a second language. Similarly, no significant difference was found between speakers of different first language backgrounds.

#### **2.1.5. More Advanced Syntactic Structures Studies**

Second language researchers have also been attracted by the study of English advanced syntactic structures such as relative clauses and the relation between the proximity of nouns and verbs in understanding the meaning of a

sentence.

Chomsky (1969) set out to investigate how young children would interpret pronominal reference, and whether children acquiring English as a first language would simply take into account the surface closeness of nouns or verbs in their processing of complex structures. In other words, Chomsky hypothesised that they applied what she called the Minimal Distance Principle. Her contention was that young children might interpret the noun nearest to the verbs as the subject of the sentence, as in

1. After *he* got the candy, Mickey left. (Would the children identify the correct pronoun referent?)
2. Pluto thinks *he* knows everything (idem)
3. *He* found that Mickey won the race (Would the children recognise the non-identity of the pronoun?)
4. The doll is *easy/eager* to see. (The doll sees...)
5. Bozo *asked / promised* Mickey to sing. (Mickey sings)
6. *Ask / tell* Bozo what to eat. (Bozo eats ...)

Van Mettre (1972) found that Spanish-English bilinguals followed an 'orderly sequence of the structures' which was parallel to that of first language learners. Cook (1973) described the acquisition of the *easy / eager* distinction for adult learners in terms of developmental stages similar to those found in Chomsky's first language learners. D'Anglejan and Tucker (1975) found that adult second language learners, especially in early stages, drew heavily on surface proximity rather than syntax to process sentences. Thus, Syngle (1973) argued that mother tongue background does not affect both adults and children's comprehension of complex structures.

However, relative clause studies revealed that, for this particular area, the



learner's mother tongue influences not only the production but also the comprehension of sentences. Schachter (1974) analysed free compositions from a group of learners whose mother tongue had a relative clause structure similar to English, or not. Schachter found that students whose relative clause structure was similar to English (e.g. Arab students of English) use as many relative clauses as native speakers do; but they make a relatively high proportion of errors in relation to certain subtleties of English relative clauses.

On the other hand, students from language backgrounds which have quite a different relative clause structure (e.g. Japanese and Chinese learners of English) produce fewer relative clauses but make fewer errors when they do produce them. Berktau's (1974) study on relative clauses production and comprehension gives further support to Schachter's findings. Berktau found that Japanese students have problems with comprehension on relative clause following subject Noun Phrases and also they (Japanese) scored lower than Spanish students. Concerning production, Berktau claims that individual learners vary greatly and that it is not possible to construct an Interlanguage continuum on basis of a study of relative clauses.

Hart and Schachter (1976), Following Schachter (1974) examined more closely student production of certain sentence constructions using written compositions from American Language Institute (ALI). The subjects came from five different language backgrounds; namely, Arabic, Chinese, Japanese, Persian, Spanish at three levels. They focused on the so-called embedded constructions, involving the following structures:

1. **Nominalisations:** *to, that, and -ing.*
2. **Relativisation:** relative-subject; relative-object; relative object of preposition; and other relatives

3. **Noun complements:** *to, that, -ing*

4. **Verb Complements:** *to that, -ing.*

The actual study involved the analysis of each composition in relation to the production, and then, the tabulation of the correct and incorrect occurrences of each structure. The incorrect responses were referred to as error types. The totals of each construction were compared in relation to levels and also to language background. The final results showed that: (1) the verb complements (especially the verb *to* complement) and relative clauses (especially the relative subject clauses) had been acquired well; (2) noun complements and nominalisations were acquired late; (3) the language background of the subjects revealed interferential difficulties. The Japanese and the Chinese, as in Schachter (1974) were found unable to produce the relatives as easily and as frequently as the Persian did.

## **2.2. Theoretical Models for Second Language Acquisition**

### **2.2.1. Krashen's Model**

#### **2.2.1.1. The Natural Order Hypothesis**

The similarity of findings in certain morpheme acquisition studies have led to the suggestion that grammatical structures are acquired in a predictable order. In other words, under certain circumstances, both formal and informal linguistic environments allow second language acquisition to take place. This view, however, does not seem to be shared by everyone.

Some studies have emphasised that adults not only increase their L<sub>2</sub> proficiency in informal environments, but they may do better than learners who spent a comparable amount of time in formal situations (Upshur, 1968; Mason,

1971; Carroll, 1967, etc.) Carroll (1967), for instance, found that students of a foreign language who spent a year's study abroad performed better than those who have never been in the country where the target language was spoken. Carroll also claimed that there was a significant correlation between a proficiency test performance and the amount of target language used at home (parental use).

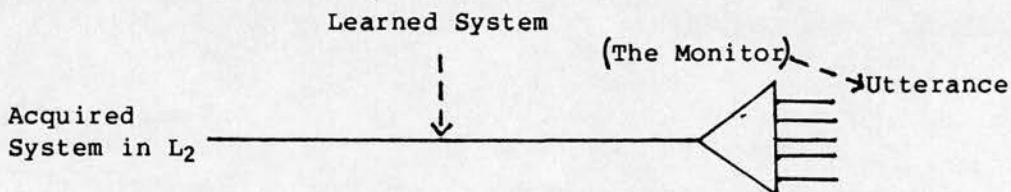
Other studies support the hypothesis that formal study means higher proficiency (Krashen, Selinker and Harnett, 1974; and Krashen and Selinker 1976). There are also claims that informal exposure increases proficiency in the second language (Krashen, Jones Zelinski and Usprich, 1978). But in Krashen et al (1976, 1978) the notion of 'exposure' is not clear. In Krashen et al., the surveys of 'years spent in an English speaking country' did not necessarily mean that the students spent their time in meaningful informal environment. For instance, Krashen et al, did not take into account how much real and sustained language their subjects used per day.

Friedlander, Jacobs, Davis and Westone (1972) emphasised that there is a difference between 'heard language' and 'intake', that is, the relevant primary linguistic data are those which the acquirer is actively involved with; thus, the total linguistic environment is less important. The results of both Upshur et al and Krashen et al, groups of studies seem to support the hypothesis that the informal environment can be efficiently utilised by the adult second language learner. Krashen et al seem to indicate that acquisition from the informal environment requires regular and intensive language use. There is no counter-evidence, Krashen adds, of the hypothesis that formal study, or its essential characteristics is significantly more efficient than informal exposure in increasing second language proficiency in adults.

### 2.2.1.2. The Monitor Hypothesis

Krashen (1976, 1978) proposes his Monitor Theory in an attempt to clarify the different roles the two linguistic environments can contribute to second language competence. Krashen makes a distinction between acquisition and learning. His main argument is that there are two way in which adult learners gain productive competence in a second language. The first is through acquisition which is an unconscious process fostered by exposure to and interaction with linguistic input in the natural environment. The second is through learning which results in the 'conscious representation of pedagogical rules'. Monitoring takes place when the learner uses his conscious knowledge of the target language rules to modify his performance in the second language. Three conditions are suggested for successful Monitor use: (1) the performer must have the time; (2) the performer must be focused on form or correctness; and (3) the performer needs to know the rule or have a correct mental representation of the rule to apply it correctly (Krashen 1977: 154).

**An Illustration of Krashen's Monitor Theory (Krashen and Terrell, 1983)**



### 2.2.1.3. The Acquisition / Learning Distinction Hypothesis

Krashen's acquisition vs. learning distinction is probably the most controversial construct in second language competence today. This hypothesis claims that adults have two distinct ways of developing competence in a second language. The first way is via language acquisition, that is, by using

language for real communication. Language acquisition is the 'natural' way to develop linguistic ability and is a subconscious process (Krashen and Terrell, 1983). For instance, children are not necessarily aware that they are communicating.

The second way to develop competence in a second language is by learning, that is, 'knowing about' language, or having a 'formal knowledge' of a language. While acquisition is unconscious, learning is conscious. The former leads to implicit knowledge whereas the latter leads to explicit knowledge of rules. It is this conscious and explicit knowledge that serves as a Monitor, under the conditions specified above. Moreover, Krashen's (1981, 1982) contention is that language teaching in grammar-based approaches which emphasise explanations of rules and correction of errors results totally in learning, not acquisition.

Krashen's acquisition-learning distinctions may, at best, have some relevance to second language research methodology, but they raise a number of theoretical and practical problems.

It is intuitively reasonable (and even empirically testable) to argue that, when 'one has time and focuses on form', one produces utterances which are variable (both qualitatively and quantitatively). Even native speakers are said to display such variability in their speech. Keenan (1977), for instance, reports some differences in unplanned talk data, planned speech, and highly planned written language samples of native speakers are not as syntactically organised as is commonly believed. Tina Bennett, working with Keenan, has suggested that the talk of native speakers shares many features of pidgins, creoles, and second language learner talk.

Krashen's model is somehow relevant to second language research, that is, in a similar (but not identical) manner as Labov's (1969) 'Observer's paradox' is, except that Krashen is only concerned with monitored vs. unmonitored style shifting whereas Labov views variation as a range of styles. Labov's first three axioms appear to be more relevant to illustrate this difference:

- a. There are no single-style speakers. Every speaker shifts linguistic and phonetic variables as the situation and topic change.
- b. It is possible to range the styles of a speaker along a continuous dimension defined by the amount of attention paid to speech.
- c. In the vernacular style, where the minimum amount of attention is given to speech, the most regular and systematic of phonological and grammatical patterns are evidenced. Other styles tend to show more variability (Labov, 1969). Thus, at worst, the Monitor theory can be referred to as only a partial model of Second Language acquisition since monitoring, that is, the conscious application of grammatical rules is an 'either-or' option, not a continuum of styles.

Tarone (1983) has argued that an adequate model for analysing interlanguage data should be concerned with accounting for variability in interlanguage along the full range of speech performance. Thus, Tarone goes on, the underlying 'Interlanguage Capability' is conceived of as an unbroken continuum of speech styles. Similarly Hyltenstam (1977) believes that rather than the existence of two types of competence, a case can be made for the existence of one type of competence - a variable competence - the manifestation of which, although constrained with certain definable limits, varies from data type to data type.

From a theoretical point of view, that is, as a model of language competence, Krashen's acquisition/learning distinction is based on the assumption that the learner has internalised two different types of knowledge: unconscious and conscious, for productive competence (what about receptive

competence?). According to Krashen only the unconscious, implicit knowledge can initiate utterances, while it cannot be accessible to introspection. If such knowledge is not accessible to introspection, then one wonders how self-correction takes place in language behaviour by means of implicit knowledge, as this also occurs in a first language. It is not clear how unconscious knowledge is acquired in a formal context—along with conscious knowledge. As the learner's interlanguage develops how is this conscious transferred to the unconscious level so that it can initiate utterances? Or else can a once conscious rule which has been transferred to the unconscious level, initiate utterances; and then, act as a monitor at the same time? In other words, how do we know which level an utterance belongs to?

Krashen (1979) has admitted that 'it is difficult for an observer to know whether acquisition or learning or some combination of both is present in someone's utterances. At this moment we have no physiological measure that shows an acquisition - learning difference.' So, whether the acquisition - learning distinction is a real one may be a matter of faith rather than empirical evidence. But if such a distinction exists it must be only relative and therefore not a useful one.

From a practical (or pedagogic) point of view the acquisition - learning distinction seems to have little to offer too since as it has been suggested, both informal and formal environments can serve efficiently for internalising the second language system (Carroll, 1967; Krashen et al., 1976, 1978)

#### **2.2.1.4. The Input Hypothesis**

The Input Hypothesis is a fairly recent (but equally controversial) elaboration of the Monitor Theory (Krashen, 1982a).

This hypothesis states that we acquire (not learn) language by understanding input that is a little beyond our current level of acquired competence. Thus, in order to progress to the next stages acquirers need to understand input language that includes a structure that is part of the next stage. The question is, obviously, 'How can we understand language that contains structures that we have not yet acquired?' According to Krashen and Terrell (1983:32) we use context and extra-linguistic information (in other words, we use meaning) to help us acquire language.

Krashen speculates, further, and introduces the concept of *Net*. The idea behind this term is that when someone talks to you in a language you have not yet acquired completely he 'casts a net' of structure around your current level (i.e. around your 'i' level of competence). This net includes many instances of your 'i' + 1 (where  $i + 1$  is the stage immediately following 'i' along some natural order). The net is the result of a speaker using a language so that the acquirer understands what is said. Caretakers, for example, provide the net to young children by talking about 'here and now'. Teachers (and Foreigner Talk) achieve a similar goal by modifying their rhetoric while talking to non-native speakers in various ways. These modifications include: slowing down, repeating, restating, changing wh-questions into yes/no questions, etc. (Freed, 1980; Gaies, 1977; Krashen, 1981).

What makes the Input Hypothesis somewhat appealing to both teachers and L<sub>2</sub> workers is undoubtedly its emphasis on the role of 'comprehensible input' and understanding the meaning of structures as the basis for language acquisition. Paradoxically, however, Krashen's theory does not allow for learned knowledge (that is, formal instruction) to lead to acquired knowledge. In other words, the Input hypothesis is the most severe criticism against whatever goes



on in classroom (grammar) teaching.

White (1987) has criticised Krashen's Input Hypothesis on grounds that, by concentrating on meaning and context, he misses the fact that certain aspects of grammar development in the learner are largely internal driven and independent of context or meaning. Moreover, White goes on, Krashen overestimates the role and benefits of simplified input. In particular, there are circumstances where the  $L_2$  input will not be able to show the learner how to retreat from the fossilized non-target form; that is, the input hypothesis is geared to handling additions to intermediate grammars rather than losses.

The Input Hypothesis, however, remains an insightful concept, especially in view of the central role it gives to meaningful communication in second language acquisition. Nevertheless, as White (1987: 107) suggests, at certain stages in the learning process, fine-tuned grammar teaching might also be a useful source of input, a means to stimulate change and lead to a different stage in  $L_2$  development.

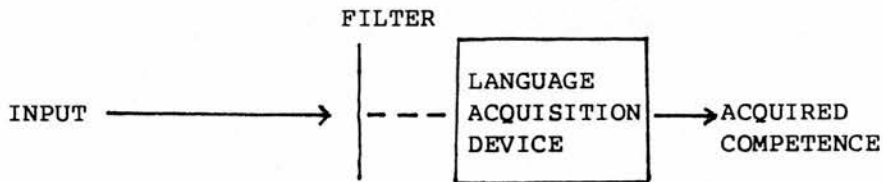
#### **2.2.1.5. The Affective Filter Hypothesis**

This Hypothesis states that attitudinal variables relating to success in  $L_2$  acquisition generally relate directly to 'acquisition' but not necessarily to 'learning'. Studies in the literature indicate that certain affective variables are related to  $L_2$  achievement. Performers with certain types of motivation (usually, but not always 'integrative') and with good self-images do better in  $L_2$  acquisition. Similarly, the best situations for language acquisition seem to be those which encourage lower anxiety levels.

Burt and Dulay (1977) have suggested that attitudinal factors may relate to

SLA in this way: performers with optimal attitudes have a lower affective filter. A lower affective filter means that the performer is more 'open' to the input. Thus, having the right attitudes may do two things: (1) to encourage the learner to try to get more input, and (2) to interact with speakers of the target language with confidence and be more receptive to the input they provide.

**Illustration of the Operation of the Affective Filter**



Summarily, the Affective filter acts to prevent input from being used for language acquisition. From this point of view, Krashen's hypothesis is of obvious interest to the present work, in as much as it points to the pedagogical relevance of social and psychological distance as discussed in Schumann's (1976, 1978) Pidginisation Hypothesis.

**2.2.2. The Acculturation Model**

Brown (1980) defines acculturation as the ability of the learner to relate and respond easily to the foreign language culture. He also identifies four stages of acculturation, which include: 1) initial excitement and euphoria; 2) culture shock, leading to feelings of estrangement and hostility towards the target culture; 3) culture stress, involving a gradual and vacillating recovery; and 4) assimilation or adaptation to the new culture. Brown argues that stage (3) is the crucial phase. Because children are less culture-bound than adults, they move through the stages of acculturation more quickly and so acquire the target language faster.

Neufeld (1978) offers a different, and somewhat more satisfactory account of the effect of age and other affective factors on SLA. He distinguishes 'primary' and 'secondary' levels of language. Primary levels include basic mastery of pronunciation, substantial functional lexical items and syntactic rules. Secondary levels include the ability to handle complex grammatical structures and stylistic varieties of the target language. Neufeld's proposal, supplemented by cognitive factors, can explain why the natural route of acquisition is not influenced by the learner's age. According to this theory, no difference in route would be observed between children and adults if the innate abilities account for the acquisition of primary levels. Adults are likely to acquire primary levels more rapidly because of their greater cognitive abilities. Children, however, will prove far more successful learners as far as the secondary levels are concerned, because they are strongly motivated to become part of the target language community and require a native-like accent to achieve this and are less affected by culture stress.

Schumann (1978) argues that acculturation, and hence SLA, is determined by the degree of social and psychological distance. According to Schumann, social distance determines whether the learning situation is 'good' or 'bad'. An example of a 'good' learning situation is when: 1) the target language and L<sub>2</sub> group view each other as socially equal; 2) the target language and L<sub>2</sub> groups are both desirous that the L<sub>2</sub> group will assimilate; 3) both the target language and L<sub>2</sub> group expect the L<sub>2</sub> group to share social facilities with the target language (i.e. there is low enclosure); 4) the L<sub>2</sub> group is small and not very cohesive; 5) the L<sub>2</sub> group's culture is congruent with that of the target language group; 6) both groups have positive attitudes to each other; and 7) the L<sub>2</sub> group envisages staying in the target language area for an extended period. When these conditions are reversed, a 'bad' learning situation obtains.

The psychological factors are affective in nature. They include: 1) language shock (i.e. the learner experiences doubt about possible confusion when using the L<sub>2</sub>); 2) culture shock (i.e. the learner experiences disorientation, stress, fear, etc. as a result of differences between his or her own culture and that of the target language community); 3) motivation; and 4) ego boundaries.

Both social and psychological distance influence SLA by determining the amount of contact with the target language that the learner experiences, and also the degree to which the learner is willing to obtain relevant input for L<sub>2</sub> development. Thus, in 'bad' learning situations, the learner will have to access to little L<sub>2</sub> input. When the psychological distance is great, the learner will fail to process the available input into intake.

Schumann's (1978) conceptualization of the acculturation phenomenon have resulted in what has become known as the 'Pidginization hypothesis'. Basically, his hypothesis is an attempt to describe the kind of learning which takes place under the circumstances of a 'bad' learning situation, involving social and psychological distance. According to Schumann 'pidginization may characterise all early SLA and . . . under conditions of social and psychological distance it persists' (1978: 110). When pidginization persists, the interlanguage fossilizes, i.e. it ceases to undergo interim revisions in the direction of the target language. Thus, Schumann concludes that: 1) early fossilization and pidginization are identical processes; and 2) continued pidginization is the result of social and psychological distance. The degree of acculturation leads to pidgin-like language: 1) by controlling the level of input that the learner receives, 2) by reflecting the function which the learner wishes to use the L<sub>2</sub> for.

### 2.2.3. The Nativization Model

Andersen (1981) proposes that SLA results from two general processes, namely, 'nativization' and 'denativization'. Nativization consists of assimilation, i.e. the learner makes the input conform to his own internalized view of what constitutes the target language. In this case, the learner builds a number of hypotheses based on the knowledge he already possesses of both his  $L_1$  and of the world. Nativization is apparent in pidginization and the early stages of first and second language acquisition. On the other hand, denativization involves accommodation, i.e. the learner is involved in adjusting his internalised system to an external norm.

In terms of strategies, the learner makes use of inferencing strategies in order to reshape his interlanguage in accordance with the input. Denativization is manifested in later first and second language acquisition, and in depidginization (i.e. the elaboration of a pidgin language through the gradual incorporation of linguistic features from an external language source).

Like the Acculturation Model, Andersen's Nativization Model explains why  $L_2$  learners often fail to succeed in reaching a native-like competence.  $L_2$  learners may be cut off from the necessary input because of social and psychological distance. Andersen's notions of 'internal' and 'external' norms indicate that the internal mechanisms play a crucial part in shaping the early or late interlanguage. However, neither Schumann nor Andersen address the issue of the differential contribution of the interaction between situation and learner.

#### 2.2.4. Accommodation Theory

Giles (1977) theory of speech accommodation is concerned with the motivation and social consequences which underlie changes in people's speech styles. A basic postulate of the theory is that people are motivated to adjust their speech style towards others. Giles proposes that the extent to which individuals shift their speech style towards, or away from the speech style of their interlocutors is a mechanism by which social approval or disapproval can be communicated. A shift in speech style towards that of another is referred to as 'convergence' whereas a shift away from the other's style of speech is called 'divergence'. As Simard et al. (1976) point out, people will reduce linguistic dissimilarities between themselves and others, that is converge, if they desire their approval and wish to integrate with them. The latter will react favourably towards those who shift towards them in speech provided they do not regard the intent of such convergence as a threat or a challenge to their own ethnic group.

Divergence or non-converging speech, however, is an important medium often used by ethnic groups as a symbolic tactic for maintaining their identity and cultural distinctiveness. One of the examples of divergence cited by Giles et al. (1977) involves when for the first time, the Arab nations issued their oil communique to the world, not in English as they did usually, but in Arabic. These nations behaved in this way in order to increase the communicative distance between them and the rest of the world in general, and their trading interlocutors in particular.

Giles et al. (1973) suggest that the strategies of convergence and divergence do not consist of simple binary socio-linguistic choices speakers

make depending on their definition of the interactive situation. Different types of convergence may be placed along a continuum of speech styles involving high, medium, or low social concessions. Giles et al. found in an analysis of speech in an inter-ethnic context that English Canadian speakers could converge towards a French Canadian listener in at least 14 different ways. For instance, some speakers would shift totally into French, others would use a mixture of both French and English but would translate certain key words and concepts into French, while still others would remain totally in English but deliberately slow down their speech rate.

According to Giles et al. (1977) and Gardner (1979), motivation is the primary determinant of  $L_2$  proficiency. The level of motivation is a reflex of perceived social distance between individual learners (i.e. the ingroup) and the target language community (i.e. the outgroup). Individuals define themselves in ethnic terms, on the basis of certain key variables, including: 1) identification with the ingroup; 2) inter-ethnic comparison; 3) perception of ethnolinguistic vitality; 4) perception of ingroup boundaries and 5) identification with other social categories. One of the most significant aspects of Giles accommodation theory is that, unlike Schumann's acculturation model, it is dynamic in nature. According to Giles, intergroup relationships change in keeping with the shifting views of identity held by each group vis-a-vis the other.

Accommodation theory also accounts for the learner's variable linguistic output as part of the group strategies in search of a positive social identity. Giles et al. (1977) point out that people continuously modify their speech with others so as to reduce or accentuate the linguistic (and hence) social differences between them depending on their perceptions of the interactive situation. Tajfel's (1977) theory of intergroup relations (combined with Giles's

theory) suggests that subordinate group members in general and L<sub>2</sub> learners in particular, have a number of possibilities open to them or strategies for linguistic differentiation. These strategies include: 1) the assimilation of the group as a whole; 2) a redefinition of previously negatively-viewed characteristics; 3) social creativity, and 4) group competition.

Members of the outgroup too have their own strategies when faced with the subordinate group's attempts to reduce the dominant group's superiority. The linguistic concomitants of these strategies often lead to either the L<sub>2</sub> speaker's use of ethnic speech markers (i.e. linguistic features which mark the ingroup membership of the speaker). On the one hand 'upward convergence' involves the accentuation of ingroup speech markers. On the other hand 'downward convergence' involves the accentuation of ethnic speech markers. Ellis (1986) suggest that in SLA progress takes place when the overall predisposition of the learner is towards downward convergence.

From the foregoing account it is clear that accommodation theory encompasses both language acquisition and variable language use. It explains why people learn a L<sub>2</sub> and takes into account variability in terms of socio-psychological attitudes in different situations.

Nevertheless, ethnic identity cannot provide a full account of variable language use (especially in a foreign learning situation). In order to specify more satisfactorily the conditions under which individuals adopt 'convergent' or 'divergent' L<sub>2</sub> speech behaviour, linguistic, psycholinguistic factors (e.g. aspects of markedness theory and monitor use) as well as social and psychological attitudes have to be taken into account.



### 2.2.5. The Universal Hypothesis and SLA

The Universal hypothesis is based on two rather different assumptions or approaches to SLA and its governing properties. Firstly, it has been suggested that there is an independent linguistic faculty, common to all humans, that is responsible for language acquisition (Wode, 1984). Dulay and Burt (1977) speculate further on this proposal by arguing that there is a 'cognitive organiser' the role of which is to apply inductive procedures on the primary linguistic data in order to create or construct the underlying IL system. In this sense, regularities of the IL system are seen as the product of this cognitive organiser rather than of an independent faculty of language.

The second assumption has the concept of 'linguistic universals' as its focal domain; that is, researchers in this area argue that certain linguistic properties influence the way in which language acquisition takes place.

The study of linguistic universals has attracted considerable interest among both L<sub>1</sub> and L<sub>2</sub> researchers over the last three decades; and it is this aspect of the universal hypothesis that will concern us in the present section (Chomsky 1965, 1980; Greenberg 1966; Comrie 1981; Rutherford 1982, 1984; Cook 1985).

Two main approaches have been identified in relation to the study of linguistic universals. Chomsky (1965, 1980, 1981) seeks to determine linguistic universals by the in-depth study of a single language. The ultimate goal within this framework is to discover the abstract principles of universal grammar applicable to all natural languages. Greenberg (1966) and Comrie (1981) represent the other approach and are mostly concerned with the identification of linguistic universals by examining several languages, with the ultimate aim of establishing what these languages have in common, i.e. 'typological universals'. We will examine the main tenets of each of these two approaches

in turn.

### 2.2.5.1. Universal Grammar

According to Chomsky (1980, 1981) the language properties inherent in the human mind make-up a 'Universal Grammar' which consists not of particular rules of a particular grammar, but of a set of general principles that apply to all grammars and that leave certain parameters open.

Universal Grammar (UG) sets the limits within which human languages can vary. One of the parameters that is open in Universal Grammar is the pro-drop parameter which is concerned, roughly, with the relationship of government between Subjects and Verbs (Cook 1985). English happens not to have pro-drop (i.e. a Subject is required for every sentence and it cannot be inverted with the verb in declarative sentences). For instance, a native speaker of English knows that the sentence *The weather is getting better* is grammatical but *\*is getting better the weather* and *\*is getting better* are not. Spanish, on the other hand, is a pro-drop language and because of this, the equivalent of these sentences would be grammatical. In other words, in Spanish 'empty' subjects can occur and inversion can take place, and is even compulsory in certain circumstances. Hence, a particular grammar amounts to a specification of the ways in which it selects from the different possibilities inherent in UG.

In Chomsky's view, language acquisition is the growth of the mental organ of language triggered by certain language experiences. Thus, the theory of Universal Grammar is largely regarded as part of biology; and to acquire language, the child need not only the Universal Grammar, but also evidence about a particular language. He needs to hear sentences of English to know how to fix the parameter for the order of Verb, Subject, and Object.

The evidence the child encounters can be positive (i.e. actual sentences of a language) or negative, both direct and indirect. Direct negative evidence consists of corrections of the child's mistakes by adults. Indirect negative evidence is provided by the non-occurrence of something in the language the child hears; that is, the fact that he never hears Subject - Object - Verb order is negative evidence that English is a Subject - Object - Verb language.

It has been suggested that first language acquisition relies chiefly on positive evidence; the child apparently receives little direct negative evidence in the form of correction of syntax (Brown and Hanlon, 1970). The few corrections that occur are mostly about dialectal or socially prescribed politeness formulae or stigmatized forms, which constitute a small area of English.

The importance of indirect negative evidence is difficult to assess since it is impossible to determine what or everything the child does not hear. Its value is hypothesized, however, on grounds that the child has already certain expectations about language that have to be fulfilled; i.e. indirect negative evidence presupposes a Universal Grammar in the child's mind.

Chomsky (1980, 1982) specifies the main tenets of Universal Grammar theory. First of all, UG is concerned with grammatical competence (i.e. the speaker's knowledge of the language) not with pragmatic competence (i.e. the ability to place language in the context of its social use, relating functions to the linguistic forms).

Secondly, in recent reformulations of UG theory (transformational) rules play a less central role than in the earlier Chomsky (1965) version. Chomsky (1982) states that rules are in effect consequences of 'principles of Universal Grammar' and of the way in which particular parameters are set. A grammar

consists of a specification of the values of parameters which may be represented as rules, but these are of secondary importance.

Thirdly, a new distinction separates 'core grammar' (i.e. those parts of the language that have grown in the child through the interaction of UG with the relevant language data) from 'peripheral grammar' (i.e. the parts outside the core, such as certain elements or forms borrowed from other languages somewhat accidentally). In terms of language acquisition, the child's mind 'prefers' not to adopt peripheral solutions, as they fall outside his pre-programmed instructions. So core grammar is less demanding than its peripheral counterpart.

Additionally, both core grammar and peripheral grammars are weighted differently in the child's mind, and are expected to be situated at the opposite ends of a continuum of markedness from core (i.e. unmarked rules) to periphery (i.e. marked rules).

Finally, Chomsky's UG theory has its strength as well as several weaknesses. The strength of this theory rests on what is known as 'the argument of the poverty of the stimulus', i.e. speakers know things they could not have learnt (thanks to certain properties of their own mind). Chomsky himself argues that the argument of the poverty of the stimulus is not peculiar to linguistics but is also valid for other areas concerned with development.

The main weakness or criticism against UG theory has been voiced in connection with its abstraction from the everyday world. More specifically, competence is separate from performance, grammatical competence from pragmatic competence, acquisition from development, core from periphery . . . all of which have relevance for actual language use from a cross-linguistic or

cross-cultural perspective.

### 2.2.5.2. Typological Universals

The second main approach to the study of linguistic universals is that of Greenberg (1966), elaborated upon by such scholars as Hawkins (1983), Comrie (1981), Keenan (1978), Thompson (1978), and Gass (1979, 1982, 1984).

The focus within this approach is on the discovery of regularities in the way that languages vary, and on the constraints and principles that underlie this variation (Hawkins 1983: 6). The kind of data used in this tradition is based on surface features which manifest themselves across a large sample of natural languages, rather than from an in-depth analysis of a single language like in the Chomskyan approach.

Three types of universals are identified. They are: 1) substantive; 2) formal and 3) implicational; and the first two are applicable to both the Chomskyan and Greenbergian approaches, whereas the third type is normally considered only within the typological tradition. Substantive universals consist of fixed features such as the distinctive phonetic features of which sounds are made, or syntactic categories such as noun, verb, subject and object. Formal universals are abstract statements about what grammatical rules are possible. For instance, much of Chomskyan linguistics has been concerned with the formulation of the formal properties or principles constraining the kind of grammar that the child can develop. Some of these are core rules and can be defined within the context of Universal Grammar. Other properties form the periphery and are best dealt with within the Greenbergian tradition. Implicational universals, on the other hand, relate the presence of one linguistic property to the presence of some other property or properties. In other words, if property X is present in a given language, then it can be assumed that

properties Y, Z . . . N are also present.

Implicational universals can be defined in terms of a cluster (i.e. the properties are related in such a way that any one of them implies the existence of the others), or, a hierarchy (i.e. the properties are ordered in such a way that the presence of one property implies the presence of all the properties higher, but not lower, in the hierarchy). Verb - Subject - Object word order and post-noun adjectives constitute an example of an implicational clustering, in the sense that if a language has, say, Subject, then it also has Verb - Object - and post-noun adjectives.

On the other hand, Comrie and Keenan's (1978) 'accessibility hierarchy' for relative clauses is one example of an implicational hierarchy. Comrie and Keenan propose an ordering for relative clauses based on the grammatical function of the relative pronoun, viz: Subject > Direct Object > Indirect Object > Oblique > Genitive etc.

This hierarchy states that if a language permits relative pronouns with a certain grammatical function (e.g. IO), then it will also permit relative clauses with all the pronouns higher (i.e. SU and DO), but not necessarily the pronouns lower in the hierarchy (i.e. Obl. and Gen.).

Comrie (1984) also distinguishes between 1) absolute and 2) statistical or universal tendencies. The former are properties which are shared by all languages and the latter, determined by considering all languages or only those of a given typology, are properties which have a statistically low to high probability of being found in at least one language.

From the foregoing discussion it can be seen that there are many ways in which linguistic universals have been investigated. In particular, examining

certain universals from a Greenbergian approach can show how these linguistic properties constrain the way in which interlanguages can develop (e.g. as a function of increased proficiency in a target language).

### 2.2.5.3. Universal Constraints on Interlanguages

The issue of how and to what extent linguistic universals contribute to interlanguage development has raised a good deal of interest among L<sub>2</sub> researchers over the last decade or so. A rapidly growing number of empirical studies have been conducted along the lines of two related but distinct theoretical questions.

The first of the theoretical questions involve the relationship between linguistic universals and 'channel capacity' (i.e. memory capacity and general cognitive abilities). According to Chomsky, first language development is influenced by both the Universal Grammar principles and non-linguistic factors such as memory and processing limitations. In other words, if these restrictions were removed, the language manifested by the child would represent pure acquisition.

Gass and Ard (1980: 445) speculate that 'patterns in first language acquisition may be much more modified by extra-linguistic exigencies than are patterns in second language acquisition.' Cook (1985) also suggests that SLA might be regarded as 'acquisition minus maturation'; and, therefore, the natural order of SLA may be closer to the true 'acquisition' order than the natural order of L<sub>1</sub> acquisition. Other scholars have suggested, however, that channel capacity (i.e. cognitive processes) still seems to play a certain role in some aspects of SLA.

According to Ellis (1982a), although L<sub>2</sub> learners display the ability to

produce fairly long utterances right from the beginning of SLA, they are still likely to evidence some features of semantic simplification in their spontaneous speech. Ellis suggests that  $L_1$  and  $L_2$  learners produce very similar kinds of utterances in the early stages (i.e. the utterances of both types of learners are telegraphic as a result of propositional and modality reduction, and other processing limitations). Cook (1975) has shown that these speech processing constraints (e.g. memory) operate in SLA as well as in  $L_1$  acquisition. However, the empirical problem remains to be resolved in order to establish which cognitive processes are transferred from  $L_1$  and which are exclusively re-activated in SLA.

The second theoretical issue concerning the role of universals in interlanguage development has to do with 'hypothesis testing', which is often regarded as one of the central processes in second language development. The arguments against hypothesis testing derive from the Universal Grammar theory, as regards the role of negative evidence. Cook (1985) points out that the  $L_2$  learner is not likely to encounter negative feedback <sup>would not</sup> and therefore, be able to disconfirm certain hypotheses which he might have induced from the available data.

Allwright (1975) and Long (1977) indicate that even in the classroom setting where sufficient negative feedback is made available, such evidence remains largely erratic and inconsistent, and seems to have only a marginal effect on the overall interlanguage development. Thus, Cook (1985) speculates that hypothesis testing in SLA is acceptable only in the sense that the learner uses positive evidence to fix the parameters of UG.

From an empirical point of view, a number of studies have focused on three main issues in connection with the effect of universals in interlanguage



development. These issues are: 1) interlanguage, like other natural languages, is subject to the constraints imposed by linguistic universals; 2) implicational universals can be used to predict the order in which properties of the L<sub>2</sub> appear in interlanguage; 3) L<sub>2</sub> learners learn less marked properties before more marked properties of the target language.

For the purpose of the present work, I will review only a few of these studies to illustrate the kind of empirical evidence which underpins the various claims about the role of linguistic universals from a typological perspective.

First of all, in the area of phonology, some insightful findings have been documented, even though the data available is undoubtedly much less than in the other aspects of interlanguage (viz syntax, semantics and pragmatics). Tarone (1980) investigated syllable structure preference in IL phonology. Her work was based on Korean, Cantonese and Portuguese learners of English. She found evidence of a preference for open (CV) syllables independent of L<sub>1</sub> background. In Tarone's view, this preference is due to the fact that the 'syllable type CV belongs to the grammar of all languages' (Clements and Keyser 1983: 28). The learners in Tarone's study simplified the syllable structure by means of either deletion or epenthesis (i.e. the insertion of one or more sounds into a word to make it conform to the phonological pattern of the borrowing language, as in Latin *schola* (school) > Spanish *escuela*).

A study also based on syllable structure by Sato (1983) presented data from two Vietnamese children learning English. However, the results were not compatible with Tarone's, since the two children did not show a tendency to use open syllables. Although Sato was not able to provide an explanation for this discrepancy, it has been suggested that at times L<sub>1</sub> or L<sub>2</sub> influences are strong enough to counter the influence of universals (Gass and Ard 1984).

Eckman (1981, 1984) investigated the issue of violations of language universals in non-native languages. In considering data from phonology and syntax, Eckman found that when violations of universal constraints do occur, they can be accounted for on the basis of the contact situation; i.e. there are explanations based on either the  $L_1$  or  $L_2$  to justify the violation. Eckman (1981) presents data to this effect from the acquisition of final voice contrasts by Spanish and Mandarin learners of English. English has a voice contrast in word final position, whereas neither Spanish nor Mandarin does. In terms of universals, it is to be noted that no language has a rule of 'schwa paragoge' (i.e. a rule which inserts a schwa following a word final voiced obstruent). Thus, if universal constraints are followed, speakers of these languages are predicted to resolve the final voiced contrast problem in some way or other than by adding a schwa word finally. In the  $L_2$  speech of the Spanish speakers in Eckman's sample, subjects resolved the problem by devoicing in word final position. Mandarin speakers learning English, however, did use a rule of schwa paragoge, even though it violates a universal prohibition against such a rule.

Secondly, in the area of syntax, Dryer (1980) presents a universal hierarchy of sentential complements. He suggests that cognitive principles constitute the basis for this hierarchy. According to this hierarchy, complements are most likely to occur in clause final position and least likely in clause internal position: clause final > clause initial > clause internal.

Given the basis of the universal, it is expected that  $L_2$  learners will behave in accordance with its principles. Frawley (1981) found that the hierarchy proposed by Dryer was also followed by his  $L_2$  learners. They used more clause final sentential complements than clause initial complements than clause internal ones.

Schmidt (1980) investigated coordinate structures in order to determine whether the types of coordination used by  $L_2$  learners followed universal constraints. Schmidt's data came from German, Finnish, Mandarin, Japanese and Arabic  $L_1$  speakers learning English. Some of the sentences used to investigate the coordination phenomena included:

1. John plays the violin and Mary the piano.
2. John typed and Mary mailed the letter.
- 3.\*John the violin and Mary plays the piano.

Schmidt found that while some of the coordinated structures that the learners produced were not similar to either  $L_1$  or  $L_2$ ; however, they did not violate universal constraints. Additionally, Schmidt found that the most difficult deletion proved to be (2) above; arguably, because object deletion is more marked than subject or verb deletion.

Evidence from the syntactic domain, but with a pedagogic perspective in mind, comes from Gass (1982) showing that within an instructional framework, generalizability occurred in accordance with the principle of the Accessibility Hierarchy (Keenan and Comrie 1977). In her study Gass found that students who were given instruction on a universally more difficult position were able to generalize their knowledge to an easier position, whereas the same generalizability did not occur from the easy positions to more difficult ones.

Hyltenstam (1982) found that the deletion / retention of pronominal copies in relative clauses by  $L_2$  learners could be accounted for in terms of the principles of the Accessibility Hierarchy. Tarallo and Myhill (1982) also found similar results in their investigation of relative clauses by English  $L_1$  learners of German and Portuguese (right-branching languages) and Chinese and Japanese

(left-branching languages). Higher hierarchal positions were easier to relativize than lower ones. Additionally, Rutherford (1982) provides some data showing how markedness factors can influence SLA. He suggests that the acquisitional order for wh-questions can be explained by markedness theory, i.e. from simple questions (e.g. *What's that?* or *What are those?*) to embedded questions (e.g. *I don't know what those are* or *I don't know what this is*).

In the area of semantics, Kumpf (1982) investigated tense / aspect using data from untutored L<sub>2</sub> learners. Kumpf found that the tense / aspect system of those learners did not correspond to the system of either the L<sub>1</sub> or the L<sub>2</sub>. Kumpf speculated that in such cases, the IL system reflects the capacity of humans to create unique form to function relationships; and more importantly, these newly-created systems will correspond to universal principles of natural languages. For instance, Kumpf observed that one of her subjects had created an aspectual system whereby a morphological device was used for marking completed vs. non-completed action, unlike that of either the L<sub>1</sub> or L<sub>2</sub>. Kumpf explains this by suggesting that in the world languages as well as in child language, aspect is primary to tense. Additionally, Kellerman (1979) examined to what extent the notion of core meanings influences SLA, in sentences such as:

4. I broke the glass.
5. The book case broke by falling.

Kellerman (1979) found that those meanings which were closer to the 'core', that is, were more 'basic' in meaning, were more likely to be accepted in the L<sub>2</sub> than those which were furthest from the core. For example, Dutch learners of English were more likely to select (4) as acceptable in English than (5) despite the fact that both sentences are acceptable in Dutch and in English.



In a different study, Gass (1984) set out to investigate L<sub>2</sub> learner's interpretation of complex sentences, more specifically sentences in which the main verb is or behaves like *promise*, *tell*, or *ask* (i.e. similar to those utilized by Chomsky (1969, 1972) in child language development). Gass obtained data from 12 different L<sub>1</sub> backgrounds. The subjects were at four different levels of proficiency and were asked to interpret the complex sentences as a function of syntactic / semantic and pragmatic factors. In her study, Gass focused on the role relationships between first NP and the second one in the sentences, in terms of an implicational topicality hierarchy. Gass found that a HUMAN noun was more likely to be selected as a topic than an ANIMATE noun to be selected as a topic than an INANIMATE noun to be selected as a topic:

HUMAN >  
ANIMATE > INANIMATE

According to Gass (1984) this ordering is independent of syntax, and the relationship between the two noun phrases (i.e. subject and object in *The boy asked the doll to leave*) is largely determined by universal factors. Similar results are presented in Chapter Six of the present study.

All these studies show that language universals may influence how L<sub>2</sub> grammars are formed. More specifically, the studies present evidence to support the three hypotheses formulated earlier, viz. 1) universals place constraints on interlanguage; 2) interlanguage development may follow universal implicational orderings, and 3) unmarked, or less marked features may be required before marked or more marked features are internalized.

### 2.2.6. Variability Analysis Models

Variation has been established as a theory in sociolinguistics for a long time now. In the study of language in its social context, Labov (1966) defines a *Sociolinguistic variable* as one which is correlated with some non-linguistic variable of the social context; the speaker, the addressee, the audience and the setting.

In linguistic theory, however, language has traditionally been distinguished from variable performance. The primary concern of the linguist has been the discovery of the invariant structure of the linguistic knowledge of the ideal speaker-hearer of a homogeneous speech community (Chomsky, 1965).

Within the framework of generative grammar, an *Obligatory rule* operates on all input strings that satisfy its structural description, whereas an optional rule may or may not apply to a satisfactory input string. Such an *Optional rule* has been referred to as 'free' variation, that is, caused by the interference of physiological factors such as memory limitations, slips of the tongue, performance errors in the individual, etc.

However, once we are faced with language use data we realise that features of individual's performance are not that random; instead they appear to be systematic in nature. Thus the notion of 'optionality' (as expressed in a traditional grammar) becomes a spurious one; and it can, in fact, be considered a failure to capture the nature of the of the systematic 'free' variation observed on the level of every individual speaker. The fact is that the option an individual chooses is clearly not as 'free' as the linguist assumes it is. Instead, it appears to be the subject to regular constraints revealed through patterns of covariation with elements of the linguistic environment and with the non-linguistic factors such as age, class, social and psychological contexts.

But the notion of optional rule of a generative grammar type cannot reveal how these elements of the structural description of a rule favour or constrain its operation. And yet, linguistic performance is characterised by a good deal of covariation or systematic dependence of rule operation frequency on both details of the structural description and non-language factors.

Studies of speech behaviour have demonstrated that the potentialities contained in the abstract optional rules can be reproduced in a fairly principled pattern in a given speaker and in a given community. It is therefore possible to argue that performance phenomena are typically variable in nature and are also part and parcel of the linguistic competence.

#### **2.2.6.1. Labov's Quantitative Paradigm**

Two opposing views have been expressed in relation to variation and its place in linguistic competence. The first view attempts to account for variation from the social, dialectal, situational and temporal standpoint. This approach assumes the existence of several linguistic systems, each occupying a certain position along a continuum within the speech community (polylectal grammars). On the individual level this approach attempts to reduce the number of environments in which variation can occur. More specifically, scholars in this approach have objected to Labov's emphasis on the probabilistic aspects of the grammar (Bickerton, 1971). In some ways, these scholars' attitudes can be regarded as an effort to down play the importance of variation and explain it with reference to the generative grammar model. This approach has been best represented by such authors as Elliot, Legun and Thompson, (1969); DeCamp, (1971); Bickerton (1971, 1973, 1975).

The second conflicting view is exemplified by Labov (1969, 1970, 1972); Bailey (1971); and Cedergren and Sankoff (1974). Most of the work done within

this perspective has been based on Labov's (1969) Quantitative Paradigm of change proposed in his study of the English copula in Non-Standard dialects. The main tenet of the paradigm is to incorporate systematic variation into linguistic description and theory by extending the notion of a rule of grammar to that of a variable rule. In this case, the predicted relative frequency of a rule to apply becomes an integral part of its structural description. The quantitative paradigm, therefore has wide implications for stylistics, sociolinguistics, dialectology, diachronic linguistics, etc. Thus, variability can be regarded as a function not only of the presence or absence of linguistic elements but also can be constrained by extra linguistic factors.

#### 2.2.6.2. The Additive Model

Linguistic rules are conventionally conceived of in a generative grammar as having the following general form:

$$X \rightarrow Y / A - B$$

e.g. Contraction:  $\partial \rightarrow \emptyset / \#\# \left[ \_ , + T \right] C \begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \#\#$

(Labov, 1972: 93)

where X is always rewritten as Y in the stated environment but is never rewritten as Y otherwise. This is a categorial instruction - the only type of rule which is permitted in any of the traditional approaches to formal grammar.

When one is faced with the fact of variation, that is, the rule does not always apply, then it is possible to say that the rule is optional. In other words, it may or may not be applied at the discretion of the speaker. We can represent such optionality by writing parentheses around the right hand member of the rule:



$$X \rightarrow (Y) / A - B$$

e.g. Contraction:  $\emptyset \rightarrow (\Phi) / \# \# [ -, +T ] C_1^0 \# \#$

However, this notation should not be interpreted as corresponding in meaning to the conventional label 'optional' if it is intended to account for facts of systematic variation referred to earlier. In Labov's sense, a variable rule of this type should achieve a higher level of accountability than unconstrained free variation. Variable rules depend on a more general principle of accountability which is normally required in the analysis of linguistic behaviour.

In this respect, Labov (1972: 94) suggest that 'any variable form (a member of a set of alternative ways of saying the same thing) should be reported with the proportion of cases in which the form did occur in the relevant environment, compared to the total number of cases in which the form did occur in the relevant environment, compared to the total number of cases in which it might have occurred'.

The first step in the formal recognition of the principle of accountability is to associate with each variable rule a specific quantity and which denotes the probability of the rule operating. This probability predicts the ratio of cases or frequency with which the rule would apply to the total population of utterances in which the rule would apply in the specified environment if it were a categorical rule of the type  $X \rightarrow Y / A-B$ . Labov has suggested that all environments for a given rule (e.g. a preceding vowel favours copula contraction, regardless of the grammatical constituent following the copula; the presence of a following verb also favours contraction, whether the preceding segment is a vowel or consonant) are governed by a fixed set of feature effects, combining in a highly predictable way. In other words, a given feature

is somewhat dependent on the presence of other aspects of the environment for a rule to apply.

This statement has led to what is known as the Additive model, which assumes that the rule probability  $p$  in a given environment is simply the sum of a number of quantities, one for each relevant feature in the environment, thus,

$$p = p_0 + \alpha_i + \alpha_j +$$

where  $p_0$  is an 'input probability' common to all environments and is a fixed number which enters into the formula if and only if feature  $i$  is present in the environment. In the additive model the effect of a given feature depends only on its presence and not on the other aspects of the environment.

The problem with the additive model is that although it is applicable to a wide class of rules, it cannot be successfully used in many others. The reason for this is that variable rule application probabilities are numbers between 0 and 1; and therefore, a general model for VARIABLE RULES should not be capable of predicting application probabilities outside the 0 and 1 interval. For some rules, however, an ordinary additive model will predict such values, especially when application frequencies are very different in different environments; or when there are large numbers of different environments. Another (but weaker) criticism against the additive model is that it assumes a numerical computation facility as part of competence.

### **2.2.6.3. The Multiplicative Model**

Cedergren and Sankoff (1974) have suggested a more satisfactory solution to the problem raised by the additive model for analysing variable rules: the Multiplicative model. Their model is basically similar to Labov's additive model,

in terms of its main properties; that is, each factor is associated with a fixed effect. But in the multiplicative model the effects are not added but multiplied together to arrive at the value pertinent to a given environment.

One of the problems of the multiplicative model is to decide, for a given rule, whether it is the application probabilities or non-application probabilities which obey the multiplicative law (i.e. since a product will never give negative values). This in fact will vary from one type of data to another.

The following is an illustration of multiplicative non-application probabilities. If  $p$  is the symbol for application probabilities, then  $1-p$  is the probability that the rule does not apply; and the model is the probability that the rule does not apply. Then the model is summarised by

$$(1-p) = (1-p_0) (1-p_i) (1-p_j) \dots$$

where  $p_0$  is, as in the additive model, an input probability common to all environments; and  $p_i$  can be considered the probability contribution of feature  $i$  so that the factor  $(1-p_i)$  is present or absent from the formula depending on whether or not feature  $i$  is present or absent from the environment. Of course  $p_i$  are all between 0 and 1.

In the multiplicative model if, in a given case, a factor  $p_i$  is absent or equal to 0, then it will have no effect on  $P$ , or the probability of the rule. But if a factor  $p_i = 1$ , then it will interfere with the effect of the other product or  $p$ , and the rule will apply categorically.

Obviously the first of these probability factors which favours the rule is the input probability  $p_0$ . In the absence of all other constraints  $p = 1 - p_0$  and  $p_0$  will have the same properties as  $p$ . If  $p_0 = 0$ , the rule will not apply at all, if  $p_0$

= 1 the rule will apply categorically.

For a variable rule,  $0 < p_0 < 1$ . This value must vary if the rule is involved, for instance, in the process of linguistic change. It will, thus, be a function of the speaker, or group whose language is governed by the rule. The variable input is also governed by such factors as contextual style, socioeconomic class, sex and ethnic group, etc.

A general formula of a variable rule would then be formulated as,

$$p = 1 - (1-p_0) (1-p_i) (1-p_j) \dots (1-p_n)$$

For example the probability of application of 'I'm' contraction rule would be summarised as follows,

$$p = 1 - \underset{1}{(1-p_0)} \underset{2}{(1-p(\text{pro-}))} \underset{3}{(1-p(-\text{nas}))} \underset{4}{(1-p(-\text{Fut}))} \underset{5}{(1-p(\text{vb}))}$$

where  $p(-\text{nas}) = 1$ . Given sentences of the type e.g. *I am your brother, I am Killer 'Diller, I am gonna go*, the third factor in the above formula will be  $1 - 1 = 0$  and the rule will always apply. If however, we have sentences of the type, e.g. *He is here, She is my sister*,  $p(-\text{nas}) = 0$  and this factor will be  $1 - 0 = 1$ , i.e. will have no effect on the probability of the rule.

Cedergren and Sankoff( 1973, 1974) re-analysed Labov's data (1969) to illustrate the power of their probabilistic model to predict the output of variable rules. They were able to predict accurately both the contraction and deletion values found by Labov. They applied their procedure to a number of other variable rules including the pharyngealisation of Panamanian Spanish /R/ and the deletion of *que* in Montreal French. Their trials have confirmed to a large

extent the hypothesis that the variable constraints are independent.

One of the advantages of the multiplicative model is that it tends itself to a simple interpretation as to the nature of the probabilistic component of linguistic competence – a notion which was implicit already in Labov's discovery of the independence of feature contribution. Another advantage offered by the multiplicative model is the use of the notion of probabilities and predicted frequencies rather than 'statistics' or actual observed frequencies' or proportions'. Statistics, frequencies, estimates, etc. cannot be predicted with 100% accuracy; and they vary somehow between performances of an identical experiment (Cedergren and Sankoff, 1974). Probabilities, on the other hand, are fixed numbers and are not subject to random variation.

#### **2.2.6.4. The Logistic Model**

Since their inception in Labov's (1969) and Sankoff's et al (1974) work, the Probabilistic Models have undergone further developments. The maximum likelihood method has been developed into what has been known as the Logistic Model, which is an improvement on the computational properties of the Multiplicative Models (more information on this model can be found in Sankoff (1975)). Sankoff (1975) has essentially proposed the 'VARBRUL 2 and 3 programs' to implement an algorithm for estimating the parameters of the logistic model. This version has been well documented and has recently been in general use for binary data ( Labov and Labov, 1977; Sankoff and Thibault, 1977; Jones, 1975; Lindsey, 1975; Naro and Lemle, 1976,1977; Lefebvre, 1975; Rousseau and Sankoff, 1978; Sankoff and Laberge, 1978; Fasold, 1978).

### 2.2.6.5. The Variable Competence Model

The variability models presented so far have been based on Labov's quantitative paradigm of change in order to analyse the probabilities of Variable Rule application, from a sociolinguistic perspective. The following model, however, approaches variability from a different angle (Tarone 1982, 1983; Ellis 1984a; Bialystok, 1982). Basically the variable competence model emphasizes the relationship between language acquisition and language use. According to Ellis (1984a) the main tenets of the variable competence model include: 1) the 'product' of language use, i.e. a continuum of discourse types ranging from unplanned to planned; and, 2) the 'process' of language use, i.e. in terms of the distinction between linguistic knowledge (or rules) and control over this knowledge (or procedures).

Tarone (1983) provides an account of variability as the product of language use, depending on the situational contexts. According to Tarone, L<sub>2</sub> speaker's competence can be seen as involving a Capability Continuum of styles, ranging from vernacular (i.e. when the learner is not paying attention to his speech) and careful style (e.g. when the learner is required to make grammaticality judgments). Lorna Dickerson (1975) provides evidence of contextual variability according to linguistic context. Her subjects were 10 Japanese speakers learning English in the U.S.A., and data were collected using a variety of tasks (viz. free speaking, reading of dialogues, and reading of word lists). L. Dickerson found that: 1) the subject's production of the target sound /z/ was influenced by the consonant / vowel phonetic environment; and 2) the consistent use of /z/ showed progress over time. The study also evidenced that for all subjects, style shifting was systematic and closely related to verbal task.

In a similar experiment type, Wayne Dickerson (1976) explored the feasibility

of referring to variability analysis as a model for SLA. W. Dickerson's main concern was to investigate in what way language learning and language change in the individual non-native speaker is similar to that in the community of native speakers. W. Dickerson (1976) found that sound change in a target language happens gradually through a Wave Mechanism which can be captured statistically by a variable rule. For instance, a word-class having a single phonetic shape /dʷ/ as in *about, mouse, south*, etc. is decomposed by environment in an orderly way and reconstructed at the end of the change with a different shape, e.g. /əw/.

Bialystok (1982), on the other hand, provides an account of the learner's variable control of his rule system, in terms of the dual distinction between automatic / non-automatic and analytic / non-analytic. According to Bialystok, knowledge that can be retrieved easily and quickly is automatic, whereas knowledge that takes time and effort to retrieve is non-automatic. Secondly, the analysed-unanalysed continuum means that the learner possesses a propositional mental representation which makes clear the structure of the knowledge.

Bialystok further speculates that unanalysed knowledge is the general form in which we know most things, since we are not usually aware of the way in which our knowledge is structured. Additionally, both the automatic / non-automatic and the analysed / non-analysed distinctions represent continua rather than dichotomies.

Ellis (1984a) proposes that procedures for actualizing knowledge are of two types: 1) primary processes (i.e. which are responsible for engaging in unplanned discourse and draw on unanalysed / automatic knowledge), and 2) secondary processes (i.e. which are called upon in planned discourse and draw

on analysed knowledge). According to Ellis, L<sub>2</sub> development occurs as a result of: 1) the acquisition of new L<sub>2</sub> rules through the participation in various discourse modes (i.e. application of procedural knowledge); and 2) the activation of L<sub>2</sub> rules which already exist in either non-automatic/unanalysed form so that they can be used in unplanned discourse.

In this section, I have discussed various models of variability analysis. These are: 1) Labov's Probabilistic models, including the additive model, the multiplicative model and the logistic model; 2) the variable competence framework including Tarone's (1983) Interlanguage Capability Continuum, Ellis (1984a, 1986) and Bialystok's (1982) models of variable competence. The application of variability analysis models to L<sub>2</sub> speech data can enable us to understand how second language learners gain control of certain features of the target language.

Nevertheless, one of the problems often encountered when using Variability Analysis Models is the fact of grouping data from different subjects. Thus, it is often difficult to consider each subject individually. Implicational scale analysis is a procedure which can help to resolve such difficulties.

### **2.2.7. Implicational Scale Model**

The implicational scale technique, also called Salogram Analysis, is a well-known procedure to psychologists (Guttman, 1944; Torgerson, 1957). Implicational analysis, as used in linguistics, is both a device for displaying variable linguistic data in ways which reveal underlying systematicity in the data, and a theoretical explanatory model (Andersen, 1978). For example, it can be used to display the correlation between the presence or absence of the copula in the environments 'pre-noun', 'pre-adj', 'pre-loc', etc. as in Table 2.1, where A-C are the context and 1-6 the subjects involved.



**TABLE 2.1 The Implicational Scale Model (adapted from Andersen 1978)**

	A	B	C
1	0	0	0
2	25	0	0
3	50	25	0
4	75	50	25
5	100	75	50
6	100	100	100

If there is a significant portion of the subjects for whom the implicational order may not hold, such a result will be clear from the implicational table and calculation of the 'coefficient of reproducibility' (Guttman, 1944).

The coefficient of reproducibility is the result of dividing the total number of errors by the total number of responses, and it varies from 0 to 1.

According to Nie et al (1975: 533), a general guideline to the interpretation of this measure is that a coefficient of reproducibility higher than .90 is considered to indicate a valid scale. This procedure, thus reveals individuals who do not conform to the implicational order. These individuals may, then, be singled out for case studies as Guttman (1944) suggests.

The statistical package for the social sciences, SPSS (Nie et al, 1975) provides some useful measures in addition to the coefficient of reproducibility, such as the 'minimum marginal reproducibility'. This measure represents the minimum coefficient of reproducibility that could have occurred for the scale given the cutting points used and the proportion of respondents passing and failing each of the items (Hatch and Farhady, 1982: 178-181). It is calculated by summing marginals for each item and dividing this sum by the total number of

responses. The difference between the coefficient of reproducibility and the minimum marginal reproducibility indicates the extent to which the former is due to response patterns rather than the inherent cumulative interrelation of variables used. The difference is called the percent improvement. The final measure is obtained by dividing the percent improvement by the difference between 1 and the minimum marginal reproducibility. The denominator represents the largest value that the percent improvement may attain, and the resulting ratio is called the coefficient of scalability which also varies from 0 to 1; and should exceed .60 if the scale is truly uni-dimensional and cumulative.

In other words, Guttman scales must first be uni-dimensional, that is, the component items must all measure movement towards or away from the same single underlying object or continuum. Second, Guttman scales must be cumulative, that is, the component items can be ordered by degree of difficulty; and respondents who reply positively to a difficult item will always respond positively to less difficult items and vice-versa.

Some examples of applications of the implicational analysis are the use of certain socially marked non-standard forms of English, the description of post-creole speech continuum (DeCamp, 1971) the use of certain standard English forms by speakers of an English-based creole undergoing decreolisation (Bickerton, 1975), the environments in which a particular variable rule operates (Fasold, 1975; Bailey, 1973) and grammatical judgements for the application of a rule in various environments (Elliot, Legun and Thompson, 1969).

The implicational scale technique has also been widely applied to data on the acquisition and use of a second language, where the attributes may be, for example, the acquisition of particular rules, the correct use of a particular rule,

in different linguistic environments (Platt, 1979), evidence of first language transfer or, the correct use of a series of grammatical morphemes (Andersen, 1978; Hyltenstam, 1977; Simukoko, 1982; Borland, 1983; Pavesi, 1986).

### **2.3. Relevance of the Review of Literature to the Present Study**

Most studies in the literature point to the fact that second language learning is a systematic and rule-governed process. The problem of systematicity vs. variation has been a matter of ongoing controversy in modern linguistic research. Thus, the recognition of the learner's language variation as systematic rather than random can be seen as a major step since it has made it possible in recent years, to study interlanguage as a separate system in its own right. Abundant evidence from the so-called 'Developmental errors' has been referred to as an adequate explanatory procedure in describing this typically idiosyncratic system. Similarly, work in variability studies has helped to exhibit that the observable variation in second language acquisition is constrained in systematic ways.

Although various theoretical models have been proposed to explain certain aspects of second language acquisition, a comprehensive theory for second language acquisition comparable to those which have been suggested for first language is still lacking. It is with such an idea in mind that the present investigation has been undertaken, that is on a rather exploratory basis. The review of literature too has been carried out to reflect this concern.

The main objective of this work is to investigate the development of the learner's interlanguage in a classroom learning situation. The present study will concentrate on the use of definite and indefinite reference and infinitival complementation. Since, to my knowledge, no systematic investigation has

been done in this area of the IL of Zairean learners of English, the choice of the structures to be studied has been based on my experience and intuitions as a teacher of English for many years, in Zaire.

On the other hand, it has been felt that many insights can be gained by referring to certain theoretical models developed over the last decade or so, in explaining the difficulties encountered by adult second language learners.

Referring back to both the Acculturation Model and Accommodation theory, it is hypothesized that, in the Zairean situation, the learner's social group can be characterised as being somewhere between a 'good' (i.e. most favourable) and 'bad' (i.e. most adverse) learning situation. Thus, the interlanguage development is supposed to rely more on psychological factors (such as, the learner's desire to learn, motivation) rather than on social attitudes. It is further assumed in the present work that social and psychological distance may play a less important role in a L<sub>2</sub> classroom setting than in a naturalistic environment.

#### **2.4. Summary and Conclusion**

Studies in second language acquisition over the last two decades have moved considerably away from those in the early sixties. In the latter period, second language research was almost inseparable with how to develop the best teaching methods and techniques. Similarly, language learning was conceived of along the lines of 'Behaviourism'. Researchers were convinced that the language learning process was one of habit formation. Thus, second language teaching methodology was strongly teacher centred on the one hand, and on the other hand, the second language learning process was not properly seen in the wider context of man's general ability to acquire a language.

The studies under review have revealed a significant change in both

emphasis and scope in the late 70's and the 80's. First, language learning pedagogy has shifted from a 'teacher-centred' to a 'learner-centred' approach. This shift in focus has been mainly due to research findings obtained from studies of child language development. Then, more and more researchers have come to realise that the learning process is not one of a 'stimulus-response' behaviouristic type (i.e. habit formation); instead, second language learning appears to involve a creative process in which the learner constructs his grammar. The learner does so following the linguistic principles of Universal Grammar, knowledge of  $L_1$  and  $L_2$  rules, together with knowledge of the world.

Secondly, there has been a growing agreement on the fact that strong claims of the  $L_1 = L_2$  hypothesis are largely unwarranted. First and second language learning are not identical, that is, the learners concerned do not appear to produce the same types of structures at any stage of development. A weaker formulation of the hypothesis seems more acceptable, that is, since first and second language utterances show strong similarities, one can speculate that similar basic mechanisms and strategies may be in operation, in both processes, for dealing with linguistic data. The discovery of such mechanisms and strategies are central to current  $L_2$  inquiry.

It has also been pointed out that mother tongue and previous linguistic knowledge play a certain role in the learning process. But this role remains highly controversial, in terms of both its kind and degree. Such a role is even more problematic to specify in a multilingual setting than in a really bilingual one.

Finally, it is hoped that the present work will contribute some understanding of the ways in which  $L_2$  language learning takes place in rather restricted environmental conditions.

## CHAPTER 3

### SOME THEORETICAL CONSIDERATIONS ON THE ASPECTS OF IL UNDER INVESTIGATION AND FORM-FUNCTION APPROACH

#### 3.1. Introduction

The present work sets out to focus on the use of definite and indefinite reference and infinitival complementation as well as other relevant features of the learners' IL which may manifest themselves through their performance. The choice of these two aspects of the IL has not been based on any established criteria or previous systematic studies. To my knowledge such studies are not available at all, as far as the Zairean learners are concerned. The main motivating factor, however, derives from my own experience as a teacher of English to Zairean speakers of French. Concerning the use of definite and indefinite reference, for example, I have become aware that the mastery of this particular aspect poses serious problems to these Zairean learners. In fact, the problem in this area is not whether the forms of the article *the, a, 0* are not being used at all, rather the difficulty seems to lie in the question whether their functional distinction has been mastered. Studies in  $L_1$  and  $L_2$  naturalistic environments show that, at first, there are few articles both indefinite and definite; and then, they blossom through a number of identifiable stages. At one stage, they are oversupplied (that is, in contexts where they are both required and not required). At later stages, a finer distinction is established between definite and indefinite, and between definite and 0 articles. It is this further differentiation stage which appears to cause intractable problems for the Zairean learners – presumably because of the influence of French which does not have a similar distinction. In other words, the learner's problem at

the later stage of acquisition is how to 'unlearn' the rule of the definite article (that is, to identify the appropriate contexts for its use). Pindi's (1982) data obtained from students of Economics at Kinshasa University seem to substantiate the observation regarding the difficulty facing Zairean learners of English. By looking at these students' compositions, he found that errors on articles are pervasive. Here are some typical examples (the topic was 'Discuss the Economic Systems of the World'): e.g.

- 1.\* They (systems) are essentially *the* capitalism and *the* communism.
2. \**The* communism is characterised by *the* central planning.

As for the choice of the investigation of aspects of complementation, this is an area which has been widely recognised in the literature as representing formidable difficulties for learners of English, from a cross-linguistic or cross-cultural standpoint (Gass 1979; Chomsky 1972; and Kachru, 1982). We look at each of these areas in turn.

### 3.2. A Syntactic Description of the Article System

#### 3.2.1. The Definite Article *the*

From a syntactic point of view, three types of uses of the definite article have been proposed: (1) a non-linguistic anaphoric description (i.e. based on deixis), (2) a definite description with relative clause, and (3) an anaphoric description.

There seems to be a good deal of controversy as to which of these three types is more fundamental to definite descriptions. Sorensen (1959) and Vendler (1968), for example, suggest that all uses of the definite article should

be described in relation to relative clauses in one way or another. Others have suggested that it is important to consider that anaphora is not only present but more basic in all three, and that it is a relatively superficial matter whether the referent happens to be in the same sentence. Stockwell, Schachter and Partee (1973: 77) argue that within the framework of a sentence grammar (as opposed to a discourse grammar) it appears preferable to leave the interpretation of *the*, in the following examples to the semantic component:

**Type 1: Non-linguistically anaphoric**

1. The key is on the door.
2. The wind is very strong today.
3. Did you wind the clock?

**Type 2: Definite Description with relative clause**

4. The student who borrowed that book wants to return it next week.
5. The new bishop is very popular.

**Type 3: Anaphoric Description**

6. There was a big demonstration in Los Angeles yesterday morning . . .  
In the evening *the* demonstration had ended.
7. A man and a woman were chatting on TV-AM two hours ago and *the*  
woman was complimenting *the* man.

Kuroda (1966) provides a good discussion of the possibility of the transformational derivation of definite articles. He proposes the following T-rule which definitises the constituent DET (with argument in favour of anaphora):



T: N<sub>1</sub> - X - DET - N<sub>2</sub> = N<sub>1</sub> - X - THAT - N<sub>2</sub>  
 1    2    3    4            1    2    -   4

Condition: N<sub>1</sub> = N<sub>2</sub> (co-referential)

Kuroda uses such examples as:

8. Someone called *a* boy to the telephone while *the* boy was talking to a pretty girl.
9. While *a* boy was talking to a pretty girl, someone called *the* but to the telephone.
10. Someone called *the* boy to the telephone while a boy was talking to a pretty girl.
11. While *the* boy was talking to a pretty girl, someone called *a* boy to the telephone.

Baker (1966) adds that *the* is inserted transformationally when an underlying existential sentence is embedded with DET. Thus,

12. ART & There was a woman that John rescued & woman recalled the circumstances of the accident.
13. The woman that John rescued recalled the circumstances of the accident.

Baker's argument is that anaphoric *the* as in (15) arises from the same source as in (12) by the deletion of the relative clause. For example:

14. There was a woman that John rescued.
15. The woman recalled the circumstances of the accident.

But under Baker's account, definitisation seems to involve (among other things) much more embedding than mere co-referentially (see Kuroda above) would normally require. Stockwell et al. (1973) concluded that the definite article

usually indicates co-extensiveness with a particular set of NP s. In other words, the definitisation assumes the co-extensiveness of NP<sub>2</sub> with NP<sub>1</sub>, whether within a sentence (as in , e.g. 6, 7) or at an extra-sentential level (as in, e.g. 1, 2, 3); where the definite article occurs with a relative clause, then the latter defines the set of NP<sub>s</sub> (as in, e.g. 4, 5).

In both English and French, the definite article is obligatory when it is: (a) accompanying superlatives; (b) accompanying other quantifiers such as *same*, *only*, *next*, which require a unique noun; and (c) in certain idioms.

16. The (\*a) best way to get home

La (\*une) meilleure facon d'arriver a la maison

17. The (\*a) same day.

Le (\*un) meme jour.

18. To kick the (\*a) bucket. (= to die).

casser la (\*une) pipe (= to die).

But as suggested earlier, it is the omission of the definite article (i.e. generic article) which causes great difficulties. Let us return, then, to the generic definite constructions.

### 3.2.2. The Generic Article (*the*, *a*, *0*)

The term 'generic' has been used to refer to a variety of grammatical constructs. Jespersen (1933) speaks of a generic person, which virtually comprises all persons on its surface realisation, by *one*, *his*, *he*, *himself*, *you* and *we*.

1. One always finds himself embarassed when he is in a situation which highlights his stupidity.

2. You can never tell about such things.

3. We live to learn.

Jespersen (1914) also uses this term to distinguish between generic and non-generic present tense, in such examples as:

4. He is ill. vs 5. None of the brave deserves fame.

This term is also used (in Jespersen op cit. b) to refer to some 'generic restrictive relatives' which occur with personal and demonstrative pronouns:

6. He that fights and runs away may live to fight another day.

Finally, Jespersen (1933) uses the terms 'generic number' and 'generic article'. It is this latter use which will concern us here. It is, in fact, in respect to this particular use of the article that our Zairean learners experience the most difficulty. Apart from the discrepancy pointed out between English and French, the major cause of difficulty seems to be the fact that the generic article involves a complex combination of definite as well as indefinite articles. One implication of this is that both the source and use of the generic article are far from clear. Concerning its characteristics, Jespersen (1933: 212-214) notes that 'an assertion may be made to apply to a whole species or class, explicitly, by the use of *every, any all* or implicitly by certain combinations of definite/ indefinite article with singular/ plural nouns.

i. No article /0/, singular: used with mass nouns, man and woman

e.g.:

7. Oil is lighter than water.

8. Capitalism differs from Communism in many ways.

ii. Indefinite article, singular (Jespersen, 1933: 213: 'it may be considered a weaker *any*) e.g.:

9. A banana tree cannot grow in Iceland.

iii. Definite article singular e.g.:

10. The lion is the king of the jungle.

iv. No article /0/, with plural nouns, e.g.:

11. Owls have large eyes.

v. Definite article with plural nouns, e.g.:

12. The Chinese drink a lot of tea.

The last usage, however, (that is generic *the* with plurals) is not widespread and some native speakers may not find it acceptable, as in:

13. The elephants are huge animals

14. The owls have large eyes.

In sum, the surface forms of the generic article are *a*, *the*, and *0*.

Postal (1966) has pointed out that generics operate syntactically like definite in some respects. Thus, only definite and generics can occur in the following.

15. Big as *the* giant was, he couldn't lift the suitcase.

16. Strong as gorillas are, they cannot overpower Mr. T-team.

17.\*Big as *a* giant was, he could not lift the suitcase.

Futhermore, generics can be pronominalised by personal (i.e. definite) pronouns (cf Wolfe, 1967). This is a very questionable statement if one considers examples (18-21) below. As was mentioned earlier, regarding the functional

description of the definite article, deixis was conceived of a prerequisite, or more basic than anaphora. In other words, if generics can be pronominalised, this would imply that sentences (20-21) are grammatical ( i.e. as generics).

18. A dog is a carnivore, but *it* also eats vegetables.

19. Milk is nutritious, but some children don't like *it*.

20.\*Milk and eggs are both nutritious but some children don't like *the milk*.

21.\*Cigarettes are more toxic than cigars, but most people still prefer *the cigarettes*.

Stockwell et al. (1973) have suggested that, presumably, definitisation does take place in sentences like 20-21, but that the article, being a generic definite, is then, realised as 0, so that the surface forms derived from 20-21 are simply the following (22-33):

22. Milk and eggs are both nutritious but some children don't like milk.

23. Cigarettes are more toxic than cigars, but most people still prefer cigarettes.

But then, Stockwell et al.'s suggestion leaves unaccounted for the fact that *the* is also a possible generic article. The above analysis highlights the complexity involved in the source of the generic article. We shall not follow this line of argument any further.

### 3.2.3. The Indefinite Article *a, an*

According to practically all analyses available, the indefinite article *a, an* is a surface form derived from the deep structure numerals 'one'. Perlmutter (1968) has posited a fundamental dichotomy between definite and indefinite which is based on their having different origins. In his analysis, *the* and *a* are entirely independent of each other in the base. Perlmutter suggests an elaborate list of

contexts which *a* and *one* have in common. One such context is that they are in complementary distribution. Others indicate the contexts in which they occur but the definite article does not. Five indicate contexts in which neither *a* nor *one* occurs but the definite article does. The others are contexts in which neither *a*, *one* (nor *the*) may occur. Perlmutter's main objective is to show that the restrictions on *a* are stated quite simply assuming that *one* underlies it. He also shows some of the rules which account for the appearance of *a* and *the*. For example, *one* is reduced to *a* when it is an unstressed proclitic; *the* is obligatorily attached to an NP which has a restricted relative clause. Stockwell et al. (1973) endorse much of Perlmutter's argument. On the other hand, Baker (1966a, b) has suggested that (a) all indefinite NPs have existential sentences as their source; (b) there is a large well-defined set of definite NPs in which the definite article is a marker of the presence of an existential sentence, in the same or previous tree, containing the same noun, etc. Sorensen (1959) and Lees (1961), too, have argued in favour of an existential source for the indefinite article. But since this particular type of article does not seem to cause difficulty for our Zairean learners of English, this controversy is of little relevance to the present study.

### 3.3. Some Theoretical Considerations on the Semantics of Articles

To begin with, let us illustrate some of the difficulties involved in the use of expressions of definite and indefinite reference. For example,

- 1a) Could you do me a favour?, a young man says to his father.  
All right, what do you want me to do?  
Find me a wife.
- 1b) Could you do me a favour?, a young man says to his father.  
All right, what do you want me to do?  
Find me the wife.
- 2a) Could you do me a favour?, a young man says to his father.  
All right, what do you want me to do?

Find me a woman.

- 2b) Could you do me a favour?, a young man says to his father.  
All right, what do you want me to do?  
Find me the woman.
- 3a) Could you do me a favour?, a young man says to his friend.  
All right, what do you want me to do?  
Find me a woman.
- 3b) Could you do me favour?, a young man says to his friend.  
All right, what do you want me to do?  
Find me the woman.
- 4a) Could you do me a favour?, a young man says to his wife.  
All right, what do you want me to do?  
Find me a wife.
- 4b) Could you do me a favour?, a young man says to his wife.  
All right, what do you want me to do?  
Find me the wife.

First of all, each of the above utterances appears to be grammatically correct and involves either an indefinite or definite referent. However, some utterances would seem to be more or less appropriate than others, especially in relation to the particular circumstances of the speech event (e.g. speaker-hearer's assumed knowledge, social status or environment). For instance, (1a) might be a legitimate request made by a man who has lived in the urban centres (say, Kinshasa) or abroad for a long time; but, being a bit of a conservative he has just returned to his home or village with a view to find a girl to marry. Sentence (1b), however, would be inappropriate for this end. Instead, it could be appropriate in a situation where the young man is a private detective looking for a couple who are suspected of having committed some crime and are wanted by the police. The detective might show the photographs of the couple to his father and say: The husband has been found already, find me the wife, etc. Sentence (2a) sounds a bit odd; and it is hardly possible to imagine the circumstances in which it could be normally used. And, yet, both (2a) and (2b) are genuine examples given by some of my students who meant to make

the request expressed in (1a) in a free composition. Sentence (2b) however, seems similar to (1b), provided that some 'foregrounding' is undertaken by the speaker. Sentence (3a) has basically the same constraints as (2a) in relation to the social setting, although in certain unusual circumstances it might be used. But sentence (3b) would seem to convey the same message as (1b) and (2b). The last set, (4a) and (4b), might have much more serious or unpredictable consequences, especially for the speaker himself. In (4a), if the expression *wife* is non-coreferential in both occurrences, it would entail that the young man has had enough with his current interlocutor/wife. But a divorce will have to be obtained first, if the message is a serious one. Nonetheless, this utterance would be highly offensive and potentially explosive a way of telling someone you want a divorce. If *wife* is co-referential (i.e. referring to the same person in both instances) (4a) would be meaningless. (4b), on the other hand, would be ambiguous, depending on whether *wife* is co-referential or not. If the expression *wife* refers to the same entity in both occurrences, the utterance would be meaningless, since it would require that the *wife* should try and 'find' herself. Instead the *wife* might decide to find a psychoanalyst to examine the crazy young man. If *wife* is non co-referential (i.e. that two occurrences refer to two different persons), then, (4b) might be given the same interpretation as (1b).

### 3.3.1. The Notion of Specific vs Non-specific Reference

In examples (1)-(4) above, the existence of a wife/woman hinges on the distinction between *a* and *the*. In addition, the felicity or appropriateness of the requests for a wife/woman depends on the relationship between the speaker-hearer and their assumed knowledge and status.

Part of what the learner has to acquire is the distinction between existence



and non-existence. However, this may not be quite the distinction required, especially in relation to adult language users since their source language might provide them with such knowledge of the world. Nevertheless, it remains difficult even for adults to use referring expressions correctly, i.e. in terms of 'particularity' or 'specificity'.

In Standard English (SE), the use of a definite article signals the reference to a very particular member of the class. Reference to particular class members may be called 'specific' reference (Brown, 1973; Maratsos, 1971). A class member 'X' referred to specifically has distinctive properties which distinguish it from all members of the class. When saying in (2b) *Find me a woman*, the speaker had in mind not just any woman, but a particular woman.

In addition, indefinite articles account for the residual; that is, they refer to no particular member of the class or set named. In some cases, however, the reference may be directed to no member of the class at all, leading to potential breakdown in communication, but breakdown in communication is an extreme case. In many cases, though, indefinite articles imply only the notion of *one* as in e.g. 1a *Find me a wife*. Also in this category would fall the negative constructions such as *I haven't got a TV set* or *I have never come across a unicorn*. In these examples, no particular member of the class *wife*, *TV set* or *unicorn* is intended. A reference to no particular member of the class, or to no member at all may be called non-specific reference. This brings us to examine the role of generics. The generic article in SE can be marked with either *the* or *a* or *0*. It is used to mark non-specific referents which can be assumed known to the hearer as in the following examples:

- 5.) a. The swan eats fish and lays eggs in a nest.  
 b. A swan eats fish and lays eggs in a nest.  
 c. Swans eat fish and lay eggs in nests.

Each of these constructions can be used to assert a generic proposition; that is, a proposition which says something not about this or that group of swans or about any particular individual swan, but about the class of swans as such. But as Lyons (1977a: 194) warns, the status of so-called generics is philosophically controversial: so too is the correlated notion of generic, as distinct from general reference. Lyons argues that there are different kinds of generic propositions which merge into one another in such a way that it is impossible to distinguish one from the other in certain instances. It is not within the scope of the present work to elaborate further on these complexities highlighted by Lyons' view as regards the use of generics. Nevertheless, whatever the philosophical status of the generics and other propositions might be, a distinction between what is general (essential) and what is specific (or contingent) cannot be avoided in any insightful linguistic analysis of English, or any other target language for that matter.

### 3.3.2. The Notion of Shared Knowledge

The distinction established in the above section between specific and non-specific reference leaves out another important aspect of the definite - indefinite semantic system. This aspect is what has been referred to as shared knowledge (Maratsos, 1976). For instance, in Britain, if a child while playing in the living room sees the milkman and rushes to the kitchen and says to his mother *The milkman is coming*, the mother is not likely to ask such a question as *What milkman?*. Instead, she will probably go straight to the front door, open it and collect the milk, eggs, bread, etc. In Zaire, however, where dairies sell their products directly to the shops, and there are no home deliveries of milk, the mother of this child would almost certainly ask *Who's the milkman?* or *What milkman?* or even *What's a milkman?*. Of course it may be the case that a Zairean child would not produce such an utterance in the first place. But the

housewife's husband might; provided he had lived in a country where such a practice is established and provided that additional 'backgrounding' is achieved to explain who the milkman is and what kind of services he might be able to offer. A definite reference to *the* class member (X) requires not only that the speaker has a uniquely specified member X in mind, but also that the reference to *the* X be specified for his listener too. In other words, *the* X should bring to the mind of both speaker and hearer the same particular, unique member of X as referent for the expression. When this condition is not filled, the listener may be confused because he is unable to bring to mind a previously specified unique member of X to correspond to the listener's definite reference (Lyons, 1977; Maratsos, 1976).

### 3.3.3. Types of Specificity of Discourse Referents

Maratsos (1971, 1976) discusses three kinds of 'specificity' in relation to discourse referents, viz. 1) specificity by general uniqueness; 2) specificity introduced conversationally; and 3) specificity by entailment. Let us examine more closely what these three types of specificity mean.

#### 3.3.3.1. Specificity by General Uniqueness

There are a number of ways in which the speaker can ensure that his listener understands reference with the same specificity as he does. Some references are specific for everyone, even without further specification because of their general uniqueness. Such references include expressions such as *the Pope*, or *the world* or *the sun*, etc. In many cases, too, references will be specific for all because of shared knowledge of the members of the group. For instance, in a classroom, the *blackboard* generally refers unambiguously. Similarly, in the household environment, the *door* would be specific; unless we have in mind a specific door. In this case, one may have to elaborate on the earlier description,

by adding, say, the *bathroom door* or the *front door* or the *back door* etc. Specificity in these instances does not depend on the object referred to, but on the relation between the object (i.e. door) and the class membership description given by the linguistic description (i.e. *bathroom door*).

### 3.3.3.2. Specificity introduced Conversationally

Sometimes the difficulty arises from the fact that a reference specific for the speaker cannot be made specific for the interlocutor or hearer by socially shared knowledge. In the above example, speaking of the milkman in a Zairean town would leave the hearer unable to establish a referent already known to any housewife for this expression. According to Maratsos (1976), in a situation like this, the speaker must then use purely conversational devices to lend specificity to reference for his listener. What the speaker must do, in this case, is to introduce the referent with what can be called a 'specific indefinite' expression. The reason why the speaker must do this is because although the intended reference is specific for him, the listener lacks knowledge of the particular referent intended. Suppose that in example (3b) above (*Find me the woman*), the speaker was the manager of a supermarket and the interlocutor a private detective. The dialogue might have started like this: *A woman who works at cash desk no 4 has just fled with all the cash (after a false bomb-alert)*; and then the manager says *Find me the woman, please*. Once a referent in a discourse has become established as a unique member of its class for both speaker and hearer, further references to it should be definite ones. These referents are referred to by Karttunen (1968) as 'discourse referents', that is, a referent that is said to be referred to specifically in the discourse for both speaker and hearer.

### 3.3.3.3. Specificity by Entailment

One of the most difficult types of specificity is the case of entailment. In theory, the kind of specificity involved in a conversation about absent referents may seem to differ substantially from the perceptual specificity provided in a conversation about physically present referents. However, according to Karttunen (1968) the two kinds of situations are closely linked. Karttunen discusses this case of entailment which provides a relatively clear conceptual bridge between conversationally induced referents and physically unique ones. The pragmatics of entailment as discussed by Karttunen appears to follow from the fact that, quite often, simply mentioning some referents or situations necessarily entails the existence of other, immediately specified referents. For instance, by mentioning that we checked into a hotel, it would not be necessary for us to tell the listener that there was a receptionist, manager, or porters etc. in that hotel. In other words, no introductory 'specific indefinite' expressions would be required for the listener to establish these persons as 'discourse referents'; since any hotel has a receptionist, manager etc. these referents would be properly definite.

Thus, Karttunen (1968) proposes that discourse referents can be created or prepared in conversations without the use of overt verbal introduction and treated referentially, much like those in physically present contexts. In addition, the rules of definite and indefinite reference only apply similarly through different kinds of discourse. However, specific indefinite expressions are often necessary to introduce a referent to the listener when no other means can achieve this end.

### 3.3.4. A Few Problems in Definite and Indefinite Reference

The following is merely an attempt to point out some of the difficulties involved in the use or function of articles in everyday speech. Without going into detail we will briefly outline five particular areas of difficulty that we think any insightful treatment of the article system must address or at least keep in mind. These problem areas include:

1. The ambiguity of some indefinite expressions.
2. A problem with entailment.
3. Non-specific reference to particular referents.
4. Generic and indefinite expressions.
5. Definite Expressions with relative clauses. <sup>1</sup>

#### 3.3.4.1. The Ambiguity of Some Indefinite Expressions

As we saw earlier, indefinite articles can be ambiguous since sometimes they are either completely non-specific in reference or specific for the speaker but not for the listener, as in:

- 6.) *Find me a teacher.*

If the speaker is looking for a particular teacher (i.e. a reference specific for himself but not his listener) then he can go on and say: *This teacher is called Tony, he comes from Canada, and wears glasses* etc. This colloquial use of the demonstrative *this* can help disambiguate the above example. In this instance, *this* is dissimilar to the normal demonstrative which is used to illustrate (point to) things that are physically close to both speaker and hearer. In addition, in example (6), the speaker may have no particular teacher in mind, in which case

---

<sup>1</sup> For a wider discussion of these and related issues, see Maratsos 1971, 1976; Bickerton and Odo 1976; Bickerton 1981; Karttunen 1968a, b; Lyons 1977a, b; Huebner 1983

we are dealing with a non-specific indefinite reference.

### 3.3.4.2. A Problem with Entailment

Karttunen (1968), who introduced the concept of entailment, pointed out that there is not a straightforward definitional relation between the use of definite articles and the referential context that enables their use. To illustrate this point, Karttunen discusses the following example:

- 7.) I was driving on the freeway the other day when suddenly *the engine* began to make a funny noise. I stopped *the car* and when I opened *the hood*, I saw *the radiator* was boiling.

In this example, it is the fact that driving entails a car which in turn entails a particular engine, hood and radiator that allows definite reference to be made as regards the phrases (in italics) above. But Karttunen points out that not all cars have radiators; e.g. Volkswagens and electrical cars do not. Yet a speaker knowing this could still say *the radiator*; presumably, because most cars do have radiators and those that do have just one. On the other hand, even if this revised formulation of entailment was taken as the norm, it may still not allow any definite reference to be made. Let us consider another example which involves less entailment, but in which definite reference had been made.

- 8.) I was driving down the freeway when suddenly *the parachute* opened.

Since, as we know, currently cars do not possess parachutes, one is bound to wonder what makes *the parachute* a definite reference phrase. This, according to Karttunen, suggests that 'entailment' may only need to be plausible rather than necessary. This point could be made clearer by the following examples:

- 9.) We were depressed on Prince Charles' wedding day.  
The video recorder was not working.  
(\* . . .A video recorder was not working)

- 10.) My wife and I wanted to take a trip to the highland but  
 the helicopter had a propeller failure.  
 (\* . . . a helicopter had a propeller failure)

In example (9), *the video recorder* is specific definite reference, although not all households in Britain had video recorders at the time of the wedding. But in (10), the situation is more compounded. The listener may feel surprised at first, and ask, *You have a helicopter?*. In such unusual circumstances, Karttunen proposes that the speaker should overtly prepare the discourse referent. He might first say, for instance, *We have a helicopter. We often use it to go to the country; but last weekend it was not working* - unless the listener is among those few ones who do have helicopters themselves.

According to Karttunen, then, the real principle at stake is a purely psychological one: the speaker should not violate the listener's range of reasonable expectations. To this extent, reference becomes less successful as far as providing for a unique reference becomes more difficult because of the implausibility of the necessary induction.

#### 3.3.4.3. Non-specific Reference to Particular Referents

Maratsos( 1976) points out the apparently contradictory situation in which the speaker is speaking about a particular class member and may be able to properly define just the class member he is referring to; and yet, he makes use of indefinite reference to it. Let us consider some examples.

- 11.) I took out a girl from King's Buildings.  
 12.) How did you like taking out a girl from King's Buildings?  
 13.) How did you like taking out the girl from King's Buildings?

The meaning involved in (11) becomes, indeed, clearer when we ask the question, as in (12) or (13). In (13), the definite expression suggests that the



speaker is interested in the girl herself; either as open-minded, outgoing, pretty, etc. or big-headed, moody, clumsy etc. However, if the question in (12) is asked, it is likely that the speaker is interested in the girl just as one token of the general class of girls from the King's Building campus. In other words, the speaker is making indefinite non-specific reference (but to an assumed known class of referents). Because of this possibility of making either specific or non-specific reference to a particular class member, utterances like (11) (*I took out a girl from K.B.*) can take either specific or non-specific meaning.

In addition, the speaker may intend the non-referential use if he just wants to enumerate what university sites he has been dating girls from. Thus, he would say *Last week I dated a girl from the King's Buildings area; the week before I had lunch with a girl from the George square area; and last term I climbed Arthur's Seat (mountain) (with a girl from the Old College)*. In this case, we have a non-referential (non-specific) use of *a*.

#### **3.3.4.4. Generic and Indefinite Expressions**

As already stated, generic statements seem to characterize general properties of classes, not particular class members. We also noted Lyons' (1977) view that philosophical status of the generic articles and the notion of generic itself do not seem to be directly related to the general principles of reference, as discussed from a semantic point of view. Thus, there is no more we shall do than remind ourselves of this difficulty inherent in the concept of generic use of articles.

#### **3.3.4.5. Definite Expressions with Relative Clauses**

To illustrate the kind of problem involved in the role of relatives and definite reference, consider the following examples:

14.) I keep thinking about the documentary film that I watched on the TV last weekend.

15.) I keep thinking about a documentary film that I watched on the TV last weekend.

Example (14) illustrates the fact that sometimes the speaker may use a definite expression not necessarily known to the listener before, or entailed by previous context. In other words, the listener may not know that I watched a documentary film and still find no difficulty in interpreting this statement. In principle then, what the relative clause does is introduce a specific context for the reference so that the listener can identify the referent as a unique member.

In (15), the indefinite article form adds another dimension, although it is almost interchangeable with the *the*. According to Smith (1964) statement (14) implies more than (15) that I saw only one documentary film, but no direct equivalence can be found between the use of the definite and indefinite forms and the uniqueness of the referent in both cases.

### **3.3.5. Definite and Indefinite Reference in Second Language Acquisition**

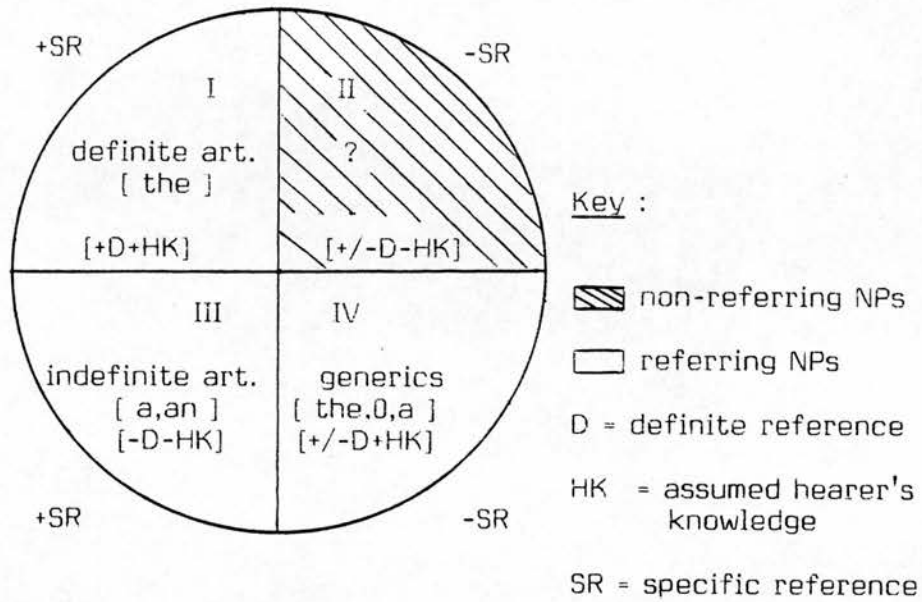
The foregoing discussion of the semantics of the articles suggests that the use of definite and indefinite reference must involve the  $L_2$  learner's internalizing an abstract sub-system, sensitive to discourse variables such as specificity and particularity. The speaker must ensure that the hearer has proper knowledge of particular referents. All languages have numerous devices to express definite and indefinite reference. Since successful communication requires proper reference to persons (as well as reference to places and times), the  $L_2$  learner must learn specific means employed for this purpose by the target language. This motivation alone would seem to justify the investigation of aspects of reference in SLA. In addition, the devices used to express reference share some properties across many languages, while in other

respects, they differ from language to language (Perdue, 1982). Successful communication can be assured so long as the learner is able to draw on what he assumes to be identical for both source and target languages. Gradually, the learner has to increase his skills at communicating successfully by adding what is different.

An appropriate investigation of the learning process in this area cannot be limited to how differently referring expressions are used, but rather, it has to analyse how, at a given point in time, the learner tries to make optimal use of the devices he has at his disposal at that time. In addition, it can be hypothesized that the learner's ability to use the skills of handling properly the different referential devices (e.g. attending to foregrounding constraints and other discourse variables) would index his L<sub>2</sub> communicative competence. It is hoped that this line of thought would be given substantial support if it were shown that the students' ability to use articles properly had some bearing on their stage of development in English.

As far as the present work is concerned, we can now conveniently talk of definite and indefinite reference in terms of the binary semantic features of specific reference (i.e. +/- SR) and assumed hearer's knowledge (i.e. + HK). These dichotomies have in fact been resorted to by some investigators in SLA to analyse data on the acquisition of the article system, in the shape and form of what Bickerton (1981) and Huebner (1979, 1983) have referred to as a 'Semantic Wheel'. In addition, the so-called semantic wheel has been re-assessed and, consequently, substantially modified, to distinguish between referring and non-referring noun phrases on the one hand, and incorporate forms of definiteness (i.e. +/- D) on the other.

Figure 3.1. A Semantic Configuration Of the Article System



Basically, two main distinctions are made. One involves specific reference (+SR), and the other takes into account non-specific reference (-SR). While specific reference is fairly obvious and more common, non-specific reference is less clearly established, since it includes properly referring expressions (such as the generics) in which the speaker and hearer share some common knowledge, as well as non-referring, non-specific expressions, in which no common knowledge or shared knowledge is to be assumed. These distinctions summarized as

I. [- Specific Referent] [+ Definite NP]  
[-Assumed Hearer's Knowledge]

e.g.

16.) General Mobutu is the *President of Zaire*. (-SR, +D, -HK)

17.) *Whoever shot the Pope* is mad. (-SR, -D, -HK)

18.) I can't drive *a car*. (-SR, -D, -HK)

19.) We are getting old; let us have *a baby* (same as in 18)

Under example (16), it is possible to express the proposition that *General*

*Mobutu likes to read historical novels, to watch football on TV, he wears spectacles* and so on. In this case, the proposition (or phrase) *President of Zaire* is not used to refer to an individual as such, but instead to say something about him. In other words, this expression is used not referentially but attributively. Similarly, under example (17), *Whoever shot the Pope* may have been referred to in various other ways. However, the relevant grounds for uttering the proposition *is mad* might not be said to refer to the individual but to the fact of having shot the Pope itself. In this sense, this proposition too is a non-referential one; rather it is used attributively. In addition, in both (18) and (19) respectively, *car* and *baby* are non-referential. And, in fact, their existence may be only prospective in that it is not certain whether I will ever be able to drive a car, own one or have a baby at all.

II. [+ Specific Referent] [+ Definite NP]  
 [+ Assumed Hearer's Knowledge] [+ SR, +D, +HK]

a. Unique referent or conventionally assumed so. e.g.

20.) The Prime Minister spoke yesterday.

b.) Referent linguistically identifiable or already established in discourse.

21.) The new Chancellor took office last week.

c.) Referent non-linguistically present but assumed known.

22.) The key is in the door.

III. [+ Specific Referent] [- Definite NP]  
 [- Assumed Hearer's Knowledge] [+SR, -D, -HK]

a.) First mention of NP in discourse but not assumed known. e.g.

23.) I bought a car in Belgium last summer.

b.) First mention of NP in an existential phrase. e.g.

24.) There is a small box on his desk.

IV. [- Specific Referent] [+/- Definite NPs]  
 [+ Assumed Hearer's Knowledge] [- SR, +/-D, +HK]

Generic use, e.g.

- 25.) The swan eats fish and lays eggs in a nest.
- 26.) A swan eats fish and lays eggs in a nest.

From the above analysis, it can be seen that expressions of definite and indefinite reference in English are very complex. On the basis of markedness theory and certain discourse variables described above, it can be hypothesised that certain articles will be more difficult to use than others. Other factors such as the linguistic environments in which they occur may also affect the learner's performance, together with other discourse pragmatic dimensions which will not be focused on in the present work.

### **3.3.6. Studies of Reference in L<sub>1</sub> and SLA**

It is important to realise that an investigation like ours, focusing on definite and indefinite reference, deals with only one of the various ways in which reference can be established in discourse. This, in turn, invites an examination of the links that exist between definite and indefinite reference and other closely related aspects of reference, viz. pronominal reference. An attempt is made in this section to review only a few relevant studies in both child language (L<sub>1</sub>) and SLA in order to establish these links between nominal and pronominal reference. Some L<sub>1</sub> acquisition studies (Bloom, Lightbown, and Hood 1975; Nelson, 1975; Weisenburger, 1976; Solan, 1978) have investigated the issue of inter- and intra- learner variability in the use of nominal reference (or anaphora). In addition, a number of SLA researchers have examined this issue in terms of the preferential use of nouns or pronouns for the expression of reference (Felix, 1977; Lightbown, 1977; Zobl, 1984).

The studies by Bloom et al. and Nelson investigated the wider issue of differing acquisition styles or strategies. In their opinion, the selection of one

nominal category over the other is related to linguistic experience and cognitive styles in cognitive development. Nelson presented evidence showing that the pronoun reference correlated with a higher Mean Length of Utterance (MLU) and that pronouns were proportionally more frequently used in sentential constructions whereas nouns tended much more to occur in isolation. The findings of these  $L_1$  studies appeared to agree on three main points. First, those children who evidenced pronominal preference tended to use nouns in non-sentential contexts. Second, in both studies, the nominal and pronominal variation was greatly diminished by the time the subject attained an MLU of 2.0 to 2.5. Thirdly, both studies propose that the rudiments of syntax can be acquired relatively independent of lexical meaning. Solan (1978), on the other hand, argued that certain facts of the acquisition of pronouns by  $L_1$  children can be accounted for not by principles of grammar, but rather by interpretive strategies (e.g. contrastive stress, or linguistic environment) and properties of language processing (e.g. short term memory, and sentence processing).

The SLA studies (Felix, 1977; Lightbown, 1977) suggest that learners give evidence of a heavy reliance on pronouns in the beginning stages of  $L_2$  learning. According to both Felix and Lightbown, pronouns permit the learner (their subjects were children) to bridge the gap between functional and syntactic knowledge available through prior linguistic experience and paucity of vocabulary knowledge marking the early stages of SLA. Zobl (1984), on the other hand, argues in favour of one of the suggestions made by Nelson, Weisenburger and Solan above, viz., that a language learner operates under processing limitations and that pronouns may take up less processing space thereby facilitating syntactic encoding. The reason for this, Zobl argues, is that pronouns generally establish reference indexically. In other words, their antecedents must be recoverable from the situational discourse context. While

it need not always be the case, their referents tend to be 'given' information (i.e. topic rather than comment). In addition, according to Thrane (1980), reference can be achieved in two ways. In a shared spatio-temporal context, reference can be achieved via indexical means; that is, 'situational reference'. The second way in which reference can be established is through descriptive naming which Thrane calls 'categorical reference' (i.e. involving the specific/non-specific distinction discussed earlier). In the latter case, an entity is identified through its membership in a category or class.

Nelson is right when she speaks of the processing economy made possible by pronominal syntax; and indeed many studies in the field have pointed out this phenomenon (Givon 1976, 1984). What Nelson has in mind is a conceptual classification brought about through the avoidance of categorical reference (for which definite and indefinite articles are normally required). Not only does categorical reference involve greater semantic specificity (as we saw earlier), indeed, in lexicalizing an argument as a noun, a speaker must observe certain selectional restrictions at a much finer level of delicacy than if the argument is lexicalized as a pronoun. Because of these differences, Givon (1976) has pointed out that pronouns encode the basic generic categories of the noun universe (e.g. animacy, sex, number). Nouns, on the other hand, (i.e. definite and indefinite reference) go far beyond these basic categories in the conceptual distinctions they draw.

On this account alone it would be plausible to propose that pronominal reference may be easier than definite and indefinite reference in both  $L_1$  and  $L_2$ . In the present work, we have chosen to focus on definite and indefinite reference, and on language universals of the type discussed by Solan (1978), Huebner (1979, 1983), Givon (1976, 1984), viz., how the use of the definite and



indefinite reference is affected by interpretive strategies, semantic features of markedness theory (e.g. linguistic environments of the noun phrase), and aspects of sentence processing (i.e. an output constraint on given / new information). Following Clark and Haviland (1977), we also believe that information processing is greatly facilitated when the listener can relate new information to known antecedents. In the same vein, Givon (1984) has proposed his Principles of Quantity Universals and Word-Order Universals, suggesting that the pragmatic structure of a communication intention places restrictions on the formal elaboration of utterance structure. Both principles are necessary in order to investigate the topic-marking universals in discourse. To summarize Givon's (1984) findings, communication could not proceed without first firmly establishing the topic, and this is normally the most urgent task. Thus, the large amount of topic repetition found in  $L_1$  as well as SLA is an illustration of this situation, where the communicative stress characterizing the system manifests itself primarily as paucity of available routinized means of identifying the clausal topic. According to Givon, the main issue at stake is a general psychological principle of task performance (viz. 'attend to the most urgent task') as well as a relatively structural question of abstract 'markedness'.

Huebner (1983: 64) has suggested that there is a tendency to place shared information (i.e. topic) before the asserted information (i.e. comment), which is common across languages. In addition, he found that the presupposed - asserted word order and the agent/experiencer - verb - object word order are compatible. If Huebner's suggestion is a valid one, then, the subject NP environment would also tend to be the most favourite for the topic; whereas the object NP environment would be usually marked for comment. The prepositional phrase environment would have an unpredictable status since time and place locatives can be placed either at the beginning or at the end of

a sentence. In terms of  $L_2$  learning of the use of definite and indefinite reference, the performance on articles *the* or *a* would presumably be facilitated in the subject NP environment than in the other two environments (since topic is shared knowledge and is easier to identify). As regards the relationship between form and function of the articles, the former would appear to proceed faster than the latter, in line with the Nelson et al.'s view that syntax can be acquired earlier or independently of lexical meaning.

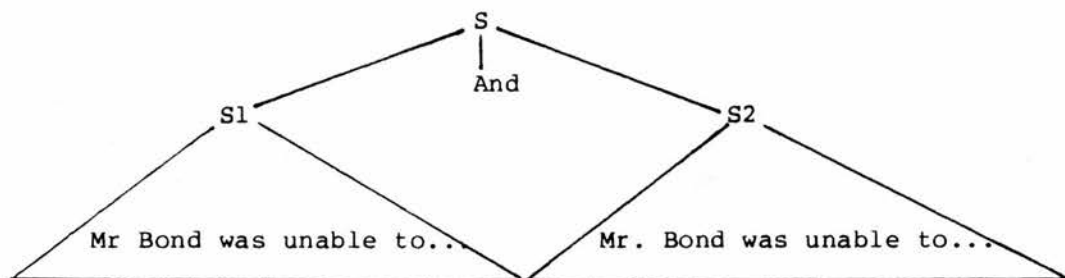
### 3.4. Some Theoretical Considerations on the Infinitival Complementation

One of the main characteristics of language is its creativity; that is, it allows us to construct and understand infinitely complex sentences within certain constraints. Sometimes a sentence is so intricate that it is not easy to specify whether it is well-formed or simply deviant. Lewis Carroll gives such an elaborate example (quoted from E. Traugott and M. Pratt, 1980: 154) Alice in Wonderland ventures that mustard is a mineral but corrects herself:

It's a vegetable. It doesn't look like one, but it is. I quite agree with you, said the Duchess; and the moral of that is - 'Be what you would seem to be' - or, if you'd like it put more simply- 'Never imagine yourself not to be otherwise than what it might appear to others that what you were or might have been was not otherwise than what you had been would have appeared to them to be otherwise'.

This type of linguistic creativity is what is called 'recursiveness'; and it is the purpose of the grammar of a language to show what (derivational) processes are involved in such constructions. Languages, in general, have two devices on which linguistic 'infinity' can be based: co-ordination and subordination. Co-ordination involves paralleling two or more structures and combining them by *and*, *but*, or *or*, under one umbrella.

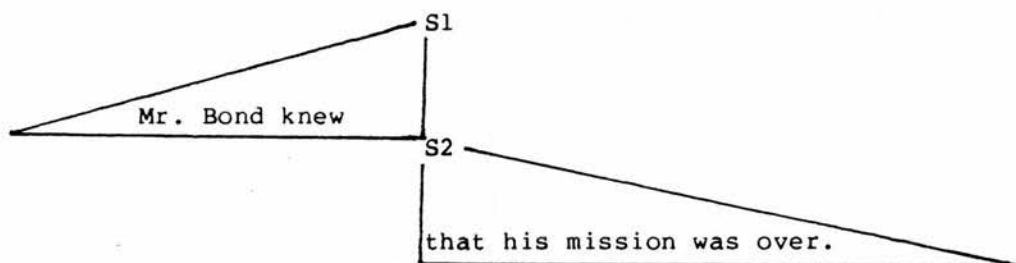
e.g. *Once in captivity, Mr Bond was unable to spy for the enemy, and further was unable to send any more flowers to his new girl friend.*



Subordination, on the other hand, imposes a hierarchic structure among sentences within a sentence, that is embedding. There are various kinds of subordination:

1. **Complementation**, e.g. *Mr. Bond knew that his mission was over.*
2. **Relativisation**, e.g. *The girl who received the flowers was a spy too.*
3. **Temporal**, e.g. *When the Duchess ate the vegetables she liked them.*
4. **Purposive**, e.g. *She decided to come late so that all the guests would notice her arrival. etc.*

Embedding basically involves a structural configuration of the following type:



The so-called complement sentences function like NPs, that is, are dominated by the node NP. For this reason, embedded sentences of this sort are often referred to as nominalisations. There are three main types of nominalisations:

- a. THAT - clauses
- b. 'S' - constructions, which are signalled by a possessive marker on the subject of the embedded sentence, and *ing* on the verb; or gerundives.
- c. FOR-TO constructions, or infinitivals.
- e.g
1. That Bill lies about it is deplorable.
  2. Bill's lying about it is deplorable.
  3. For Bill to lie about it is deplorable
  4. I asked that he should be sent.
  - 5.\*I asked his being sent.
  6. I asked for him to be sent
  - 7.\*I wanted that Bill should go.
  - 8.\*I wanted Bill's going.
  9. I wanted Bill to go.

Which type of nominalisation is used is largely determined by the verb. This information must therefore be included in the lexical entry of the verb.

In the present study, we are more concerned with infinitivals. The main objective is to investigate how the notion of co-referentiality is understood when the subject of the embedded clause ( $NP_2$ ) has undergone the Erasure Principle (Chomsky, 1965) by an equivalent NP in the matrix sentence. The motivation for the choice of the structure (that is, infinitivalisation) is that, eventually, we hope to be able to explain and understand how, in their IL relative to English, Zairean learners encode the notions of reference, in terms of definiteness, specificity and shared knowledge, and co-referentiality (in infinitival nominalisations). Thus, as in the case of reference, an account based on both form and function is being presented in the next section.

Kiparsky and Kiparsky (1970) propose that many of the differences in the form and meaning of nominalisations depend not on essentially arbitrary syntactic features but rather on semantic features in the governing items. They suggest a set of parameters, viz.: factivity vs. non-factivity, and emotivity vs. non-emotivity.

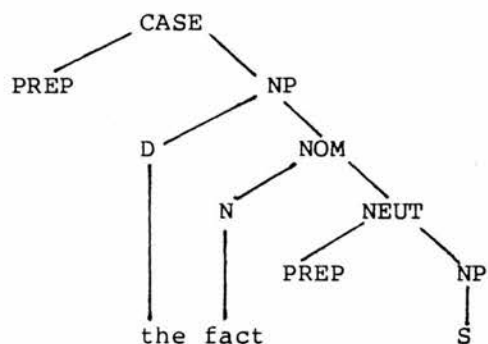
According to Kiparsky and Kiparsky, factive predicates can only occur when the speaker presupposes that the propositional object or subject of the predicate is factually true. Non-factive predicates occur when the speaker merely asserts or believes the proposition to be true, but does not presuppose its factuality. Negation is perhaps the clearest way of showing this distinction.

10. (Factive) It is odd that the door is closed.  
          It isn't odd that the door is closed.

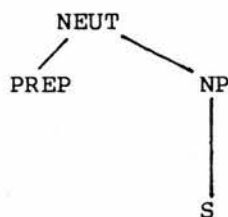
11. (Non-factive) It is likely that the door is closed  
          It isn't likely that the door is closed.

In Kiparsky and Kiparsky's (1970) analysis, factive nominalisations have the deep structure 'the fact that'; non-factive nominalisations have the deep structure 'that S':

## i. Factive



## ii. Non-factive



The second set of parameters, that is, emotive vs. non-emotive, state that predicates which express the subjective value of a proposition, rather than knowledge about its truth value, are emotive, as in e.g. *It is important for him to integrate into the community.*

Infinitival nominalisations are taken to be a secondary consequence of several distinct processes which have the effect of leaving the verb without a subject with which it can undergo agreement. In some cases the subject is marked with an oblique surface case (as when *for* is inserted with emotive predicates) or deleting it (as when it is erased by EQUI-NP-DEL). In other cases, the subject can be raised out of its own sentence as in, *He is unlikely to solve the problem.* Given these cases, the rule of infinitivalisation (i.e. TO REPLACE AUX) applies to insert the form *to* in the position of the auxiliary.

### 3.4.1. The Principle of Minimal Distance [PMD]

One important principle of transformational theory is the 'recoverability of deletion' (Chomsky, 1965: 182). The justification for this principle (i.e. the identity condition) is that an ambiguous sentence whose derivation involves a deletion could have an infinite number of sources. Thus, under an identity

condition a sentence like e.g. *She tried to leave* is assumed to contain two occurrences of the subject *she*: *She tried* + *She AUX leave*. Then, the subject of the embedded sentence is erased by the higher identical subject.

Rosenbaum (1967) developed an erasure principle which could ensure for his derivations that there could be no ambiguity as to which was the erasing NP. According to Rosenbaum, the first NP to the left is responsible for the erasure. Although this principle does work in several instances, it simply does not hold in many more.

12. They tempted John to leave early.

13. We forced John to ignore his work.

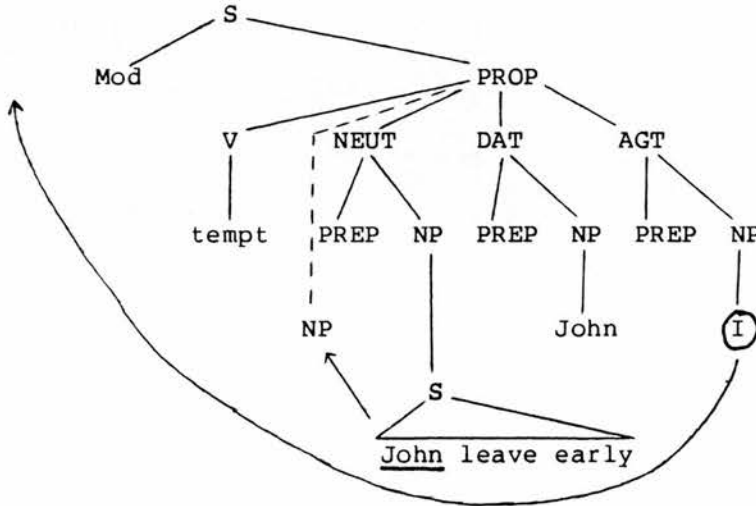
One type of clause at least where it does not work is, for example, in purpose clauses; since it would mean that *car* and *bike* be erased in the following:

14. John sold his car to buy a house.

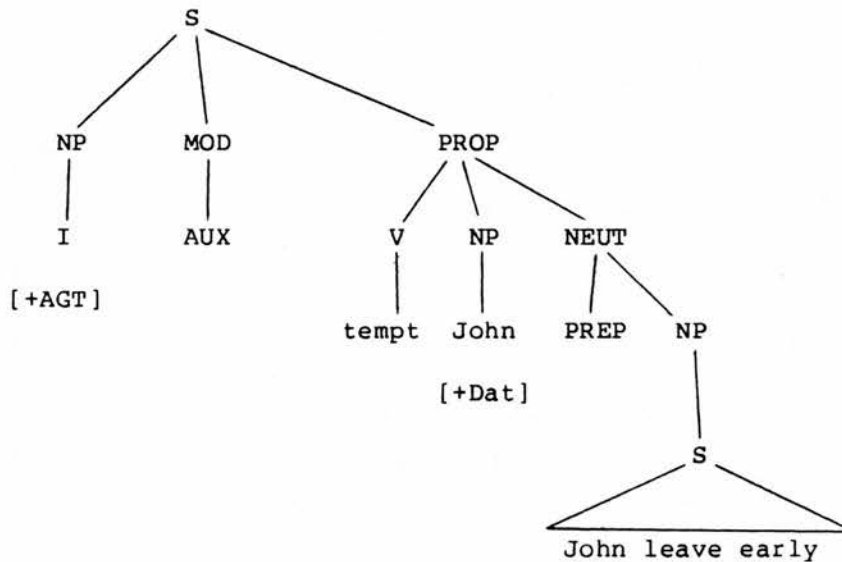
15. Mr. Tebbit took his bike to look for a job.

### An Illustration of the Principle of Minimum Distance

- (a) The structure of prior to the application of the Subject and Object raising in the early cycle:



- (b) The result of the application of the Subject and Object raising:



Rosenbaum (1967) and Chomsky (1968) have posited the so-called Principle of Minimal Distance, measured by counting the number of branches in the path connecting two nodes. This principle is intended to eliminate the problem of the 'purpose clauses', since the subject of the purpose clause is (presumably)



more distant from the matrix object than from the matrix subject. Stockwell et al. (1973) have used Fillmorean (1968) trees to show why this principle of minimal distance is rather shaky (see tree-diagrams a, b). Apparently, according to Rosenbaum's (1967) proposal, the rule of EQUI-NP-DEL, which deletes the subject of the embedded S (that is *John* in tree diagram a), must apply very early before the Case placement rule which moves the appropriate NP (that is /) into surface subject position.

Stockwell et al. (1973) stress that of necessity, the EQUI-NP-DEL rule must precede the raising of the subject of the embedded S to object of the matrix S (as shown by the dotted branch in tree A), in order to allow normal reflexivisation to take place, as in :

16. John believes himself to be intelligent., but block it in

17. \*John wanted himself to work hard.

### 3.4.2. The Dative / Agent Principle

Stockwell et al. (1973) have questioned the accuracy of the way in which the distance between the subject of the embedded-S and the erasing subject is determined. They have argued that in the example *They tempted John to leave early* the distance of the correct erasing NP, the dative NP, is identical to the distance of the other matrix NP, that is, the Agent - since the EQUI-NP-DEL must precede the case placement and reflexivisation rules. In view of this dilemma, Stockwell et al. have proposed that the erasing NP is identified by the case node dominating it. By doing so, they have replaced the principle of Minimal distance by the principle that an *identical* dative has erasure priority over an identical agent. But, as Stockwell et al. themselves recognise, neither

Rosenbaum's (1976) and Chomsky's (1968) Minimal Distance Principle nor their own Dative/Agent principle can describe nominalisations satisfactorily (largely because they do not pay sufficient attention to semantic criteria). The difficulty with their analysis can be illustrated in the following examples:

18. ? He promised us to leave at once.

The problem with such a sentence, for example, is whether it should be interpreted as *He promised us that he would leave at once* or *\*We were promised to leave at once*; or else a combination of *He promised us that he would leave at once* and *He promised to leave at once*.

The foregoing remarks have tried to show, very briefly, that infinitivalisations are very complex structures, and that they are the more so when we attempt to deal with them ignoring the semantic features (like  $\pm$  factive,  $\pm$  emotive, etc) of the dominating node.

Learners of English are likely to experience great difficulty in acquiring the infinitivalised constructions. The reason for this is that these constructions involve a complex combination of both syntactic processes and semantic features. The learner's task will indubitably be an extremely arduous one, particularly in the early stages (Gass, 1984). It can be hypothesised that beginners may rely more on meaning than form, e.g. on knowledge of the world, to decide which NP is the subject of the embedded S (that is, deletable). Advanced learners, on the other hand, may tend to rely on both meaning and form, but they too may fall back on meaning and pragmatics when in doubt, in interpreting infinitivals, as in (*ask, tell, promise* sentences, etc.)

19. The pupil asked the teacher to begin the lesson

20. The teacher asked the pupil to begin the lesson.

21. The pilot told the plane to land.
22. The plane told the pilot to land.
23. The father promised the children to buy a family car.
24. The key promised the door to leave.

### 3.5. Some Theoretical Considerations on the Form-Function Approach

In order to attend best to the learner's IL development, as measured by an increase in  $L_2$  proficiency and communicative effectiveness, it is important that the researcher understands its process. Process means here the background relative to the way in which the learner has created his transitional competence and the problems he may have been struggling with. Earlier work in IL tended to focus on its product, rather than its process. By focusing on the product, as exemplified in Morpheme studies, the investigators looked at a number of grammatical forms and regarded them as acquired if they were produced in environments where a native speaker would use them. There are many problems in this so-called 'Form to Form' analysis. One is that it is target language oriented. The point is that accurate suppliance of a TL form does not entail that the learner knows its function(s). Thus, the learner may produce a form without having mastered its meaning. Another problem is that the developmental patterns are often obscured by the fact that different functions or uses of the forms can easily be lumped together under one single label. For example, a learner might be thought to have acquired that COPULA because he has supplied *is* in all obligatory contexts, but never used the other forms (i.e. *are, was, were, etc.*). In such a case, the researcher may have failed to account for the role of 'avoidance' as a major strategy in IL production.

Form to form analysis, then, can only tell, perhaps what is acquired and in

what order – but not how it is acquired, i.e. IL development in a strict sense. But, apart from the theoretical shortcomings, this kind of approach to IL offers the advantage of being amenable to quantitative analysis. Thus, the findings have some potential for generalisability. In fact, the popularity of Morpheme studies was largely due to the fact that several researchers (using the same rather sophisticated statistical procedures) were able to replicate and check the claims made in relation to the 'Natural Orders' of acquisition.

As far as the present work is concerned, a detailed analysis of the sample content based on semantic features (or function) has been carried out (see sections 3.2., 3.3., 3.4.). The motivation for this was to avoid the shortcomings involved with a 'Form to Form' analysis based on the Target Language. In our case, the areas of difficulty have been identified, taking into account the learner's main background language (which is French, under most of the present analysis). Bickerton (1981) provides additional (and persuasive) argument in favour of such a stand. He argues that the acquisition of a target feature *qua* feature – a given morphological shape – and the acquisition of an accurate range of target meaning/ function are two completely different things. Bickerton (1981) suggests further, that SLA is targeted language change in which the surface forms of features are borrowed from the target language and incorporated into the semantic and syntactic structure of the acquirers' language at whatever stage this has currently reached. This preoccupation with how to trace the learner's IL variation in form-to-function (rather than form-to-form) mapping is the main characteristic underlying our sample, described in functional terms earlier in this chapter. The implication of this is that our data will be quantitatively different from those used in early PA studies, since it is (1) form-to-function based, (2) dynamic (that is, looking at variation over different points in time, at secondary and university levels), and

of course, (3) capable of quantitative analysis.

### **3.6. Summary**

In this chapter, we have provided for a broad linguistic description of the two main aspects of IL under investigation, viz. the article system and the notion of reference, as well as certain features of the infinitival complementation. However, no particular theoretical model has been advocated for describing these syntactic structures. It is common sense in SLA research to obtain as many insights as possible from whatever theoretical or practical models available within the study of language phenomena. Our sole motivation has been, then, to be able to take into account both syntactic and semantic characteristics of these subsystems of the IL developing grammar, viz. a form-function approach. In the next chapter we are going to present in detail the general experimental design and data collection method.

## CHAPTER 4

## GENERAL EXPERIMENTAL DESIGN AND DATA COLLECTION

## 4.1. Testing Instruments and Directions for Fieldwork

Studies in the literature indicate that describing the language learner's language involves a clear understanding and careful investigation of the following concepts (Labov 1972; Bickerton 1975, 1981; Corder 1967, 1981; Schumann 1978; Krashen 1976, 1981, 1982):

1. The learner's Interlanguage is variable but systematic.
2. The learner's Interlanguage is an idiosyncratic system of increasing complexity.
3. The learner's Interlanguage is constrained in both specific and principled ways: that is, depending on such variables as (a) the learner himself, (b) the setting or environment in which the learning takes place (c) certain input characteristics and/or instructional variables, and, possibly (d) his mother tongue and other previous knowledge.

Apart from our focusing on what the learner himself is doing, it is also relevant to assess the role of certain external or environmental factors in relation to Interlanguage development. Schumann (1978) for instance, has referred to 'social and psychological distance' in his attempt to explain the lack of development on the behalf of the learner's IL. Although there may be some substance in his argument, especially in a native-speaker context, the picture is much more complex, or even different, in a Foreign language classroom situation.

In the Foreign language context, there may exist a somewhat circular relationship between input characteristics and social and psychological

variables. In other words, poor or inadequate input would fail to become intake, and consequently, lead to the learner's failure to master the appropriate language forms and functions required to express his messages. More often than not, this lack of IL development (i.e. fossilisation) results in a great degree of attrition or frustration – an idea implicit in Schumann's social and psychological distance, or Burt and Dulay's (1977) affective filter. This so-called social and psychological distance may, in turn, prevent the learner from further attempting to get more input or test his hypotheses – leading to a vicious circle. Presumably, this vicious circle may go on so long as richer or more adequate input is not made available and converted into intake by the learner. If the present perspective was found to be correct, one would then suggest that the relationship between social and psychological distance and lack of IL development ought to be conceived of *not* as one of cause-to-effect, but rather as somewhat circular in kind, at least in a foreign language situation.

Looking at IL behaviour from such a standpoint has great significance for both SLA theory and practice. From a theoretical point of view, fossilization would be regarded as, potentially, leading to social and psychological distance – not the other way around. In other words, social and psychological distance is not necessarily a given fact in the learning process, and may well derive from the latter, in some cases. From a pedagogic point of view, it would appear crucial for the classroom teacher to be able to provide the learner with rich and adequate input so as to pre-empt or to tackle the well attested phenomenon of fossilisation, and so to minimize the effects of certain external factors. Providing the learner with the adequate input presupposes not just that one has a fairly good idea of what the learner's 'etat de dialecte' (Corder, 1967) is at one given point in time, but also an understanding of why it is so. A satisfactory answer to this concern would inevitably involve looking at the

learner's IL over a series of developmental stages.

For the purposes of this study, four types of tasks are envisaged:

1. One Modified Cloze Test for the article system (definite article *the*: 43 items; indefinite article *a, an*: 16 items; and *0* article: 15 items; total 74 items)
2. One sentence interpretation task, for the infinitival complements (36 sentences)
3. One picture description composition task in connection with the story behind the pictures of a traffic accident involving a lorry driver and a cyclist.
4. A questionnaire on attitudes, motivation, and desire to learn English.

In the following sections more details are given concerning these tasks and directions for implementing them.

#### **4.2. Requirements of the Tasks**

Before dealing with the different task types and their contents, it is important to keep in mind the kinds of criteria a test, or series of tasks, should meet in order to achieve its basic objectives. Harris (1969: 13) suggests that a test should be *appropriate* in terms of our objectives (i.e. the criterion of validity), *dependable* in the evidence it provides (i.e. the criterion of reliability), and *applicable* to our particular situation (i.e. the criterion of practicality).

In spite of the complexity of the issues involved in the implementation of these criteria in most testing situations, an effort has been made to meet such requirements.

#### **4.3. Task Types and Test Items**

In this section, full details are given concerning both the rationale behind the selection of these tasks. The actual items as well as the instructions for



implementing the tasks are given in Appendices II, III, IV and V.

#### **4.3.1. A Modified Cloze test for the Article System (74 items)**

Cloze tests have been used for a great variety of purposes. Initially, cloze tests as developed by Taylor (1953) were used to measure the readability of prose. The method consisted of simply deleting every *n*th word from a passage and replacing it with a blank of a standard length. When used as a measure of readability, the Cloze procedure aims at giving an estimate of the difficulty of the passage – not measuring the skills of the examinee. Since their inception, however, Cloze tests have been extended to the measurement of other skills rather than reading *per se*. Taylor himself even proposed the possibility of using the Cloze procedure to measure  $L_2$  proficiency. Other uses of the cloze procedure technique include the measurement of vocabulary usage, ability to read aloud, intelligence quotient, etc. (Oller 1973, 1978). Oller reports on the use of this technique to reveal sensitivity to even more subtle variables such as to discriminate significantly between groups of subjects who hold different opinions concerning the content of given passages of prose. Oller further mentions uses of cloze tests to detect sensitivity to extra-linguistic information provided in pictures used to illustrate textual materials. Briefly, the Cloze procedure as described here has turned out to be a remarkably stable and sensitive technique for measuring a wide range of skills or variables.

In our investigation of the use of reference, it appears appropriate to use the Cloze procedure (rather than a multiple choice test for instance) in view of the reliance of the concepts of 'definiteness' and 'specificity' on the context, both linguistic and non-linguistic. However, some modification has been made concerning the distance between the blanks so as to focus only on the articles. Thus, the task has been referred to as a 'Modified Cloze Test'. It consists of

filling in the blanks in a story where all the articles (definite, indefinite, 0) have been removed (see Appendix II) Traditionally, every fifth, sixth or seventh word is deleted from the passage. Oller (1973) points out that with native speakers, deleting words more frequently than one out of five creates a test of such difficulty that much discriminatory power is lost. On the other hand, to delete words less frequently than one in twelve does not substantially change the quality of the test. However, he goes on, a method of random deletion of a certain percentage of words from a passage of prose yields similar results, although it may be less convenient to use. Nonetheless, the parameter of distance between blanks has not been investigated systematically with non-native speakers yet.

Concerning the items themselves (or sample content), we have already pointed out (section 3.2.2) that the function of the article system is best understood with reference to such notions as specificity or shared knowledge. But we also maintained that certain syntactic features such as the Mass/Count distinction ought to be considered. So, a combination of both semantic and syntactic criteria have been applied to the sample. Three different subclasses are thus distinguished within the article system.

**1. Definite Article *the*:**

- a. Linguistically anaphoric  
e.g. We took a SOTRAZ bus at the airport and got off  
*the* bus near the Central market.
- b. Linguistically non-anaphoric  
e.g. *The* key is on the door.
- c. With an adjective preceding an NP  
e.g. *The* new bishop is very popular.

**2. Generic Article, especially the 0 article:**

- a. Based on the Mass/count distinction /0/  
e.g. A Gorilla likes *ice-cream* but does not like water.

b. Definite and Indefinite Articles singular, referring to a whole species:

- e.g. *The* lion is the King of the Jungle.  
*A* parrot cannot bark.

c. No article /0/, plural:

- e.g. Eggs are nutritious.

3. Indefinite Article: *An, a* (phonologically, it derives from the reduction of the numeral 'one')

- e.g. I have an apple and a banana in my bag.

#### 4.3.2. Sentence Interpretation for Infinitival Complements

The sentences used in this task aim at investigating to what extent the learner's proficiency level would affect his interpretation of complex sentences such as, e.g. *I wanted Bill to go* (see Appendix II). Gass (1984) has proposed that the acquisition of complex sentences cannot be understood in isolation, and that language universals affect developing grammars. Gass suggests that universals seen in the role of Topicality Hierarchy (i.e. Human > Animate > Inanimate) interact with the language specific facts of English. In substance, Gass's main argument is that second language speakers at the beginner's level are likely to depend more on semantics (or meaning) than syntax (or form) to decide which NP is the subject of the embedded sentence (i.e. the EQUI-NP-DEL). Advanced learners, on the other hand, would rely on both semantics and syntax and possibly pragmatics as well to determine the deletable NP in such sentences as for example:

- 1.) The teacher asked the student to begin the lesson.  
 (The question is: Who should begin the lesson, the teacher or the student?)
- 2.) The chicken told the house to go away. (Who should go away, the chicken or the house?)
- 3.) John promised Bill to study hard. (Who studies hard?)

The sentences used in this task contain verbs of the same type as those found in Gass (1984), viz. *ask, tell, promise*, and a few more not used by Gass - but

belonging to the same sub-class of verbs referred to as sharing the features [-factive, +emotive], in Kiparsky and Kiparsky (1970) proposal (and also in Carrell, 1984). Thus the other verbs used in this experiment are *want*, *hope*, *prefer* and *order*. All possible combinations of the features 'Human', 'Animate', 'Inanimate' are included, at least as far as the *ask / tell / promise* set is concerned. These combinations, then, give the following spectrum:

MAIN VERB	VERB OF EMBEDDED 'S'	EXAMPLES OF SENTENCES (Nos. in brackets refer to the list of sentences in Appendix III)
1.) + Human	+ Human	1.) <i>ask</i> (4,16) (i.e. a request) 2.) <i>tell</i> (1) 3.) <i>promise</i> (12, 24, 36) 4.) <i>want</i> (35) 5.) <i>hope</i> (30) 6.) <i>prefer</i> (13) 7.) <i>order</i> (22) 8.) <i>ask</i> 8, 31) (i.e. a 'question')
2.) + Human	+ Human	1.) <i>ask</i> (28) 2.) <i>tell</i> (9) 3.) <i>promise</i> (2)
3.) + Animate	+ Human	1.) <i>ask</i> (32) 2.) <i>tell</i> (17) 3.) <i>promise</i> (6)
4.) + Human	+ Inanimate	1.) <i>ask</i> (23) 2.) <i>tell</i> (10) 3.) <i>promise</i> (20)
5.) + Inanimate	+ Human	1.) <i>ask</i> (11) 2.) <i>tell</i> (25) 3.) <i>promise</i> (18)
6.) + Animate	+ Animate	1.) <i>ask</i> (19) 2.) <i>tell</i> (14) 3.) <i>promise</i> (34)
7.) + Animate	+ Inanimate	1.) <i>ask</i> (15) 2.) <i>tell</i> (26) 3.) <i>promise</i> (27)
8.) + Inanimate	+ Animate	1.) <i>ask</i> (5)

		2.) <i>tell</i> (33)
		3.) <i>promise</i> (21)
9.) + Inanimate	+ Inanimate	1.) <i>ask</i> (7)
		2.) <i>tell</i> (3)
		3.) <i>promise</i> (29)

#### 4.3.3. Picture Composition Task

This task involves the description of a car accident in four episodes as illustrated in a set of pictures (see Appendix IV). The subjects are asked not to limit themselves to the mere description of what is happening in the pictures, since this would not challenge their ability to communicate in a meaningful way. Instead, they have to tell the story behind these pictures as if they were themselves involved in the accident, i.e. from the cyclist's or driver's point of view. In other words, the subjects will be somehow obliged to 'talk with' the visual prompts, rather than 'talk about' them. The subjects are encouraged to write between ten and fifteen lines per episode, and not worry about mistakes.

#### 4.3.4. Questionnaire on Attitude and Motivation

The aim of this questionnaire is to assess the Zairean learner's attitude and motivation to study English as a Foreign Language (see Appendix V). Following Gardner and Lambert (1972) and Jakobovits (1970), Schumann (1978) suggests that a student's attitude towards Foreign Language study and culture is one of the major determinants of successful learning of that particular language.

Traditionally, L<sub>2</sub> learners are grouped into two main motivation types: (1) the integratively oriented, and (2) the instrumentally oriented learners. According to Schumann (1978), an integratively oriented learner is interested in acquiring the second language in order to meet and communicate with valued members of the target language community. A learner with an instrumental

orientation is one who has little interest in the people who speak the target language but nevertheless wants to learn it for more utilitarian reasons, e.g. in order to obtain a promotion in his occupation, or gain recognition from his own membership group.

Schumann (1978) speculates, further, that integrative motivation implies the minimal psychological distance, whereas instrumental motivation involves a whole range of psychological distance. For example, if the goal of the instrumentally oriented is mere survival (like a pupil who just wants to pass his final exams) the extent of psychological distance will be great. But if he wants to qualify for a highly competitive post, e.g. as an interpreter in a multi-national company, then, he may acquire a lot more of the target language. On the other hand, Gardner and Lambert (1972) report that lack of interest in foreign languages, ethnocentrism, etc. correlate with unsuccessful second language acquisition. The questionnaire used here has three parts; and has been adapted from Jakobovits (1970: 262-70) 'Attitude Scale' and 'Orientation Index' (see Appendix V).

Part One addresses itself to attitude. It is composed of 20 statements about English speaking people's language and culture. The statements are presented to the subjects in French to ensure that they are clearly understood. The higher the score on this part, the more favourable the learner's attitude towards English speaking people and their culture. These learners are also expected to be more successful in their achievement in the language.

Part Two deals with Motivation types (i.e. Orientation Index) and is made up of ten statements: Five for integrative (i.e. even numbers), and five for instrumental orientation (i.e. odd numbers); the degree of instrumental or integrative orientation can be determined by summing the values scored by the

learner on each of the four alternative responses; each alternative has been assigned a numerical value, ranging from (a) +2, i.e. *I agree*, (b) +1.5 i.e. *I slightly agree*, (c) 1, i.e. *I have no opinion*, to (d) 0.5, i.e. *I disagree*.

Part Three is concerned with the desire to learn English i.e. assessing the extent to what the learner himself takes certain initiatives or how eager he is to contribute positively to his learning tasks. This part is composed of ten statements too. Each statement has four alternatives, which have been assigned numerical values, i.e. ranging from +2, +1 and 0 to -1.

#### **4.4. Administration of the Tasks**

##### **4.4.1. Learner's Avoidance Strategy**

In any investigation dealing with linguistic phenomena there is always the thorny problem of first locating and contacting informants and secondly, getting them to reveal freely their actual linguistic behaviour.

The methodological issues involved in an undertaking of this kind have been expounded by Labov (1972: 181), in terms of what has come to be known as the Observer's Paradox. In his influential paper, Labov contends that styles can be ranged along a single dimension, measured by the amount of attention paid to speech. He further argues that the vernacular is the style in which the minimum attention is given to the monitoring of speech. He also suggests that 'the only way to obtain sufficiently good data on the speech of any one person is through individual, tape-recorded interview'. The Observer's Paradox itself derives from the fact that 'the aim of linguistic research. . . must be to find out how people talk when they are not being systematically observed; yet we can only obtain this data by systematic observation'. Labov (op. cit.) then proposes

that we must find ways of supplementing the formal interviews with other data. One way of achieving this would be to devise various tasks involving the use of different skills on the behalf of the subjects. When faced with describing the language learner's IL, however, the difficulties referred to above are compounded by the fact that the learner often avoids certain target language forms, until he has fully recognised them. In view of this difficulty, the investigator should design his tasks in such a way that the learner's attention is set on function rather than form - and thus minimise the negative effects of this avoidance strategy.

#### **4.4.1.1. Administrative Problems**

Apart from the methodological problems pointed out above, the investigator should also take certain administrative facts into account.

As far as the present work is concerned, most tasks are due to be carried out in the classroom/school context. Therefore the following factors have to be considered:

1. **The School System.** The School authorities and teachers co-operation is necessary, e.g. in order to encourage the students to participate positively in the experiments, or allow for certain disruptions which may derive from them.
2. **The School Time-Table.** It is important to know if, for instance, the tests can be completed within the ordinary class period, or, if extra time can be allowed in case this is not the case.
3. **Other Problems,** such as late starts, illness, absenteeism, may not be solved by the investigator. However, they have to be at least acknowledged when they do occur. In particular, public transport is a major handicap in Kinshasa and many students often report to school very late.



#### **4.4.2. Equipment**

The proposed experiments do not require the use of any sophisticated equipment or materials. The questions could be typed in advance and handed out at the beginning of the session. Normally, the answers to all the intended questions could be given using the same sheet.

In the following sections, we give an account of the various steps taken during the pilot study to assess its appropriateness of the testing instruments designed to elicit data. Then, we will describe briefly the actual procedures used to carry out the main study. Finally, an attempt is made to summarise the outcome(s).

### **4.5. The Pilot Study**

#### **4.5.1. Aim**

The Pilot study was undertaken with a view to assess the suitability of the various tasks; that is, their appropriateness in order to elicit adequate data from the subjects, within the existing administrative and education system.

#### **4.5.2. Population Sample**

The first concern here is to determine the criterion for selecting the particular school, faculty or department where the subjects would be taken from. This step was relatively easy to take as far as the tertiary level was concerned (that is, at University and Teacher-training college), but more difficult at the secondary. In Kinshasa there is only one university, and one Teacher-training college which can offer the 'Licence' degree (a university first degree).

At the time our pilot testing was due to begin (December 1984) we had to

contend with a number of administrative problems in order to locate the informants at the University of Kinshasa. We found out that for some reason, not all of the 6 faculties operating there had included English in their current timetable, or curriculum. In the Faculty of Law, for instance, English teaching had been 'temporarily' discontinued for three years. On the other hand, in the Faculty of Economics (1st and 2nd years) the course had been scheduled for the 1st semester only, ending early in February. In fact, the actual teaching was brought to an end by mid-January to allow the students to prepare for their 'out-of-session' examinations (this arrangement means that students are offered an opportunity to sit a certain number of papers, in February, to 'eliminate' subjects in which they get higher marks). In the other faculties, in January and February, it was much the same picture (that is, the Faculty of Civil Engineering, Medicine and Pharmacy). The Faculty of Sciences was the only one where the English course was still running normally at the time of my investigation. Thus, there was no choice to be made, since this was the only opportunity of obtaining a sample from the university students. English course in the faculties are organised by our own School of Languages. It was therefore, easy to make appointments to meet the students.

According to the experimental design, it was necessary to obtain a sample of 1st and 2nd undergraduates studying in the same faculty and ideally being taught by the same teacher, in order to minimise the effect of instructional variables. On this ground, the department of Mathematics and Physics was selected. The number of students involved in the Pilot study was 18 (13 first year students and 5 second year ones), all male, aged between 20-23. The pilot tests took place in February. The main experiment itself had to be conducted one month later. This category of University students will be conveniently referred to as the 'Intermediate' group.

At the secondary schools too a number of steps were needed to locate the subjects. In a sense, decision-making at this level was much more difficult; since there are so many schools to choose from. In other words, equal opportunity ought to be given to all types of schools to take part in the investigation, through some kind of randomization technique; since it was not possible to include all of them. And yet, some of the schools would appear to be better organised, equipped, or more efficient than others.

Apart from the criterion of randomization, the researcher felt somehow bound to give greater consideration to schools which could be referred to as 'good' or 'efficient' ones. Of course, the labeling of a school as 'good' as opposed to 'poor' is purely an arbitrary decision; but it is meant to imply that if a school can be proven to be 'good', it is more likely to provide for potential candidates for higher education than a 'poor' school would do. But there was nothing near a reliable measure that we could find or use to test (let alone prove) the 'goodness' of a school. We then decided that a 'good' school would mean one in which, over the last three years, an average of 75% or more of final year pupils had been successful in obtaining the 'State Diploma', that is, the school leaving certificate which gives access to Universities and Colleges of Higher Education. A list of 10 schools was established on the basis of information available at the Ministry of Education. The researcher, then, drew one name from a hat to select the school where the experiment was to be conducted. The name of the school was the Elikya Institute. The school has a good reputation in Kinshasa, for the efficiency of its administration (like most church-owned schools) and a high success rate among its pupils. It has both full-fledged primary and secondary school networks. According to the deputy Headmaster, in 1983-84, its success rate had been well above 80% for the school leaving State Examinations; and thus, many of its former pupils may

have stood a good chance of securing a place in higher education. This is a potentially significant factor since the experimenter was interested in finding a cross-section of 'equivalent' samples among Zairean students at three levels of proficiency. This group of secondary school students was the 'Elementary' one. At the secondary school, English is taught as a subject from the 3rd year up till the 6th, in all three options (namely, Literary, Physics/Mathematics, and Bio-Chemistry sections). The subjects involved in the pilot-study were drawn from the 6th form of the literary section (27 pupils). The tests themselves were administered during January '85 by the experimenter. The third group of subjects involved was that from the teacher-training college viz. the Institut Pedagogique National (I.P.N.). These are 1st and 2nd year undergrads in the English Department.

From the outset of the experimental design, this category of subjects was regarded as a control group; and thus, it was not included in the pilot study. Teacher trainees have, in effect, more opportunity to have access to adequate input, since English is the medium of instruction and subject-matter at the same time. Therefore they were credited with the ability to perform better on any of the tasks which would be found suitable for the lower levels; that is, the Intermediate and Elementary. On the other hand, if a given task were to be found unsuited (say, too difficult or too easy) for the lower proficiency levels, it would be pointless to proceed with the task at this advanced level. Data from this group was, therefore, collected only during the main study, in March 1985.

#### **4.5.3. Administration of Testing Instruments**

The actual fieldwork period lasted approximately 15 weeks involving extensive observation of the informants, through regular visits to English classes at the secondary schools, the university and the teacher's college. The

period of time leading to the end of term I was devoted to visiting a number of secondary schools, obtaining official permission to conduct the experiments during term 2, and duplicating the testing instruments. In January and February, the pilot study was carried out; and the main study in February and March. The experiments took place mostly during the morning classes and lasted approximately 90 minutes (that is, two 45-minute periods). Both the pilot study and main study were administered by the investigator himself. But at the secondary school, the pupils' teacher was sometimes present in the classroom just helping to hand out the test questionnaires. In other words, he did not interfere with the experiment in any way except that he was believed to ensure that those pupils with a 'shifty eye' did not take advantage of their more conscientious and laborious classmates.

The students were told that the researcher had come from the University of Edinburgh (in Scotland), and intended to obtain some data from them - in order to investigate and, then, suggest ways of making learning English easier for Zairean students; and improving some aspects of the Zairean education system. Therefore, the students were invited to be as co-operative as possible and were assured that the results would not be passed on to their teacher. The students seemed willing to cooperate. In addition to that, the questionnaires regarding Attitude/Motivation survey were handed out to the students after the test sessions so that they could be filled in at home to save precious classroom time.

#### **4.5.3.1. Modified Cloze Test**

This task was administered to investigate the mastery of the Article system in English. A short text was handed out to the students, after all the 'articles' had been extracted from it. The students were asked to fill in the blanks with

the article *the, a, an* or leave a blank (0) when they felt no article was needed. The instructions were printed on the answer sheet (in French) and did not seem to cause difficulty to the testees. The task, at all levels, was completed within 15 minutes. At Elikya Secondary School, the task was administered to 43 pupils (in the fifth form), and 27 in the sixth form. But the fifth form pupils' responses were discarded, on grounds that the classroom was too small, and the students were so squeezed together that it was impossible to prevent them from seeing their neighbour's answer sheets. On the other hand, the sixth form pupils were seated in well-separated rows, and the task took place in a relaxed atmosphere.

At the university, the problem of overcrowding did not arise. No further incident was recorded as far as the Modified Cloze task (pilot test) was concerned.

The outcome of the Modified Cloze Test showed that, as expected, the University students performed better on articles than the secondary school pupils did; however, this difference between the two groups was not significant, as the results of a T-test revealed ( $t = .37$ ;  $p < 0.5$ ;  $d.f = 31$ ). Table 4. 1 shows the means differences between the 2 groups.

Table 4.1

**Elementary (E) and Intermediate (I) Groups Means Differences  
on the Modified Cloze Test**

LEVEL	n	Mean	S.D.	t
E	27	43.67	10.24	
I	6	45.84	13.7	.37

This lack of significant differences between the two groups of subjects does not imply that the task was not suitable, since what the investigator was most interested in at this stage was the level of difficulty of the task (the values of 59% and 62%, for the E. and I. groups respectively, were regarded as satisfactory). As to the dichotomising power between proficiency levels, it was felt that this could be increased substantially by expanding the sample at both ends; that is, this could be achieved if the teacher-trainees (the upper end) and 5th form pupils (lower end) were included. It may well be that secondary school final year pupils and 1st year undergraduates have a good deal in common in terms of English language proficiency. However, it would be too hasty a conclusion to jump to at this stage. For the above reasons, the researcher decided to include the Modified Cloze test task, unchanged, in the main investigation.

#### 4.5.3.2. Sentence Interpretation

The Sentence Interpretation task was intended to see if these students would obey the so-called 'Minimum Distance Principle' (Rosenbaum, 1967; Chomsky, 1969), that is, if they would assign Subject status to whatever NP

most immediately precedes the verb Infinitival Constructions of the following type:

The teacher asked the student to begin the lesson.

Mary asked Louise what to eat.

It was also hypothesised that, in some cases, the students would have difficulty in interpreting the sentence if a conflict arises between syntax and semantics, and that a certain Universal Topicality Hierarchy (Comrie, 1981; Gass, 1984) would then intervene to determine the Subject of the embedded clause (i.e. EQUI-NP-DEL) as in:

The car told the soldier to stop.

The dog asked the door to come in.

The task, then, consisted of 36 sentences of the type referred to above (see Appendix III). The subjects were instructed to tick the correct subject of the infinitival construction. 27 pupils from the sixth form and 15 first year undergraduates in the Maths/Physics department were involved. The results of the test revealed that the overall performance of each of the two groups was good; that is, the test was of average difficulty (viz., 69.5% for the pupils and 72% for the undergrads).

On the other hand, a Point Bi-Serial Correlation Coefficient ( $r_{pbi}$ ) was computed to find out whether the undergraduates performed significantly better than the secondary school pupils. The formula used for the Bi-serial Correlation coefficient reads (Guilford, J.P. and B. Fruchter, 1978:309)

$$r_{pbi} = \frac{\bar{x}_p - \bar{x}_q}{st} \sqrt{pq}$$



where the symbols are defined as:

$\bar{x}_p$  = mean of x values for the higher group (i.e. for the undergrads:  $\bar{x} = 25.94$ ; s.d. = 4.48; n = 15) along the dichotomised variable (which is described here in terms of the Minimum Distance Principle).

$\bar{x}_q$  = mean of x values for the lower group (i.e. for the secondary school pupils:  $\bar{x} = 25$ ; s.d. = 7.26; n = 27)

p = proportion of cases in the higher group (i.e.  $15/42 = .357$ )

q = proportion of cases in the lower group (i.e.  $27/42 = .643$ )

st = standard deviation for the total sample in the continuously measured variable (i.e. st = 6.36)

Thus,

$$\begin{aligned} r_{pbi} &= \frac{25.94 - 25}{6.36} \sqrt{(.357) (.643)} \\ &= \frac{.94}{6.36} \sqrt{.23} \\ &= .07 \end{aligned}$$

The Point Bi-serial Correlation coefficient showed no relationship between higher proficiency level and identification of the EQUI-NP. However, since the task appeared to be of moderate difficulty, the experimenter felt confident that, like in the first task, what was needed was to broaden the sample, by including the Teacher-trainees at the one end and some fifth form pupils at the other.

The Pilot study, once again, indicated that not much English language learning had been achieved by the 1st year University students. The latter may be undergoing more pressure to concentrate on their specialist subject rather than developing their knowledge of English. On the other hand, final year school pupils are assumed to be highly stimulated to master as much knowledge as possible in all subjects, and English in particular, in order to face the highly competitive school-leaving State exams in confidence.

#### **4.5.3.3. Written Composition Tasks <see Appendix IV>**

The instruments designed to elicit this type of data were

1. A pictorial description (written composition) of a traffic accident involving a lorry and a cyclist.
2. A Free-Composition describing the subject's life over the last five years.

Both tasks were administered during the pilot study. However, only the first of the two subtests appeared to elicit appropriate data. The set of four pictures showed a clear sequence of events, enabling the student to focus on the meaning conveyed by these pictures. On the other hand, the lack of a particular topic to focus on during the second task seemed to involve the fact that the students became more concerned with grammatical accuracy; and therefore, might have been inhibited in their attempt to describe meaningful events in their life. The 'Free-Composition' task was therefore regarded as inadequate for the purposes of the present investigation.

Indeed, the written description of the traffic accident was a feasible task; and the pictures themselves were depicting scenes which are very familiar in Kinshasa (see Byrne, D. 1967: 52). On the whole, the subjects appeared to grasp

the meaning of the story behind these pictures. They had between 50–60 minutes to complete their composition. This task was retained for the Main study, whereas the Free Composition was rejected.

#### **4.5.3.4. Attitude and Motivation Survey**

The questionnaire to investigate Attitudes and Motivation consisted of three parts:

1. Attitudes towards English Language and Culture (20 items)
2. Learner's Orientation Index (10 items)
3. Desire to Learn English (10 items)

The items were written in French to ensure greater understanding (but see Appendix V for English equivalent), and handed out to the pupils or students as 'home based' activity, to save classroom time. In this respect, the subjects had plenty of time to think about their 'genuine' responses, since they were encouraged to provide truly individual answers. But this involved some drawbacks as well. For instance, a certain number of questionnaires were never returned. There is also the likelihood that some pupils may have appealed to their classmates or parents for help; the latter may have then supplied specific answers to certain questions, rather than just explain the meaning of the item. But there is little evidence that the subjects copied answers from their classmates. As for the parental influence, this would not necessarily distort the real picture, since it is somewhat reasonable to assume that to a certain extent, pupils' and parents' attitudes are closely related anyway.

## **4.6. The Main Study**

### **4.6.1. Assumptions**

The main experiment took at least 4 weeks after the Pilot study for each group. The following tasks were retained for the main study:

1. The Modified Cloze Test
2. Sentence Interpretation, and
3. The Picture Description Task

The testing instruments remained unchanged since they appeared to have an average level of difficulty, whereas the sample itself was made more sensitive by including more subjects (viz. the Advanced Group composed of teacher trainees and one lower proficiency group at the Elementary level). The pilot study revealed, among other things, that the university undergrads did not score significantly better than the final year pupils. One possible explanation is that the two groups do not differ very much; presumably because the amount of exposure falls dramatically from 5 hours per week at the secondary level to 2 hours only at the university, thus their linguistic knowledge does not increase substantially. However, it would be equally misleading to insist on obtaining impressive correlation values since this might involve a conception of language development in purely linear terms – an assumption that could not be taken seriously in view of the available evidence to the contrary in IL studies. The question would then be, how one would be able to establish the fact that there exist general developmental progressions or trends to which the learners conform. Although there does not seem to be a clear-cut answer to this question, it is conceivable to argue that learners at different levels progress or master the various IL subsystems at different stages. Having the foregoing remarks in mind, the experimenter proceeded with the main experiment, using a

much larger sample.

#### **4.6.2. Subjects**

The total number of subjects involved was 163, subdivided into three main groups, as follows:

1. The pupils' group; i.e. 5th and 6th forms at Secondary school (total Ss = 66)
2. The Undergraduates' group; i.e. 1st and 2nd year undergrads in the Faculty of Sciences, Kinshasa University (total Ss = 11)
3. The Trainee-teachers' Group; i.e. 1st and 2nd year, at the Institut National Pedagogique, English Department (total Ss = 76).

#### **4.6.3. Administration of the Tasks**

##### **4.6.3.1. At the Secondary Level**

The Main experiment took place for both the 5th and 6th form at the end of February. An agreement had been reached between the school's Director of studies and the researcher to the effect that each class could be available for 2 successive periods of 45 minutes. In other words, each class had 90 minutes to complete the 3 tasks; the 90 minute-long sessions had been found sufficient during the pilot study.

Concerning the 5th form, 41 pupils were accounted for at the end of the session. This figure includes, in fact 4 late-comers who arrived late at the end of the first 45 minutes, missing the first 2 tasks, viz. the Modified CLoze test and Sentence Interpretation. No other incidents occurred. The Modified Cloze test took 15 minutes to complete, Sentence Interpretation about 15 minutes, and the rest of the session was devoted to the Picture Description written composition (i.e. roughly 60 minutes).

As for the 6th form, 29 subjects were present, and no late arrival was

recorded. The atmosphere was relaxed and the sub-tests were completed within exactly the same time span as during the pilot study. At the end of the session, the pupils said they had experienced some difficulty regarding the written composition. They found it very hard to put meanings into words, and claimed that this was due to the fact they had to think in French and then try to express the same ideas in English. The pupils commented further that multiple choice tests (i.e. discrete items) are their favourite type rather than compositions (i.e. integrative tests). Later on, I discussed the pupils' preference with their Director of Studies. He said that pupils are in fact familiar with both types. But at that time of year, final year pupils insist on being given more tests of the Multiple Choice type because that is the one used for the State Examinations questionnaires; this does not mean they could not respond to other test types satisfactorily, he said.

#### **4.6.3.2. At the University Level**

Like at the secondary, the tests were conducted during 90 minute sessions. By the time the Main Experiment took place, the students had been divided into several small tutorial groups so much so that it was rather difficult to work out a convenient time for bringing together many subjects: 21 were available. During the time span between the Pilot study and Main experiment, the researcher became aware of the fact that the university students were concerned about ways of increasing their restricted opportunities to learn English. Some of them had been attending evening classes (at their own expense) at the University School of Languages. Most of these students, rightly believe that higher proficiency in the English language may boost their ability to master their (science-based) specialist subjects. The experimenter provided them with encouragement to carry on in the same direction. He also advised

them to gain as much background knowledge of their specialist subject as possible, along with English.

As for the test sessions themselves, they were conducted by the experimenter himself. Their English teacher did not take part; and they were presented in the following order: (1) Modified Cloze Test, (2) Sentence Interpretation, and (3) Picture description. The Attitude/Motivation Questionnaire was handed out for them to fill in at home and return to the class representative who was in charge of handing them in.

#### **4.6.3.3. At the Teacher's College**

In terms of the present study, this is the Advanced group. The sample is composed of 43 students in the first year and 33 in the second. The Modified Cloze test lasted 15 minutes, Sentence Interpretation 15 minutes, and Picture description 60 minutes; and, as usual, the Attitudes questionnaire was handed out for home study.

The testees were very co-operative. Some of them appeared to be surprised by the fact that the instructions on their answer sheets were given in French. The researcher had to tell them that this was not meant to belittle their knowledge of English, but due to the fact that the same tests had been administered to students for whom English was not a major subject for study.

It is possible that this explanation alone may have led to the students' performance being affected by the implicit competitiveness. However, it is obvious that, even if a different explanation had been given to them, these students have a good image of themselves; as it were they wanted their linguistic competence to be taken seriously. Self-esteem, as we know, is an

important and positive psychological factor in second language learning (Schumann, 1978; Krashen, 1981).

#### 4.7. Summary

In this chapter we have presented the general experimental design as well as the procedures used for pilot testing and collection of data for the main study. Before dealing with the results of the analysis of data the scope of the investigation can be defined as follows:

1. **Aim of Research** – To determine how various subsystems of the learner's Interlanguage are acquired; i.e., first from the mastery of certain linguistic forms then, towards the mastery of their appropriate functions. The aspects of IL under investigation are definite and indefinite reference, and infinitival complements.
2. **Approach** – The approach involves a quantitative analysis well as a qualitative one, i.e. a mapping of form-to-function relationship across 3 levels of proficiency. Interlanguage users generally, and Zairean learners of English in particular, are likely to produce Target-like linguistic forms before understanding fully their meaning in English. The identification of the learning process as well as the understanding of aspects of its evolvment are matters of great interest for both theoretical and empirical work in SLA.
3. **Data** – Three types of data have been gathered to investigate the learners' IL development. A Modified Cloze test has been used to obtain data on the articles. Similarly, a set of pictures showing a traffic accident is used to elicit the story behind the pictures. Moreover, a sentence interpretation task has been given to investigate the learners' understanding of complex sentences of English. A key feature of the sample content is its being analysed in semantic as well as syntactic terms (rather than in purely morpho-syntactic terms).
4. Also available is data obtained through a Attitude/Motivation survey, with a view to examine how individual learner factors can affect IL development.
5. **Particular Questions Addressed**
  - a. What is there in the data to reveal the systematicity of the IL of the learners referred to in this study?



- b. Is the amount of variability observed among the subjects a matter of random versus systematic change at different points in time?
- c. Form-to-Function mapping is in a positive relationship with level of proficiency.
- d. Distinctions such as Specific vs. Non-Specific (e.g. in the case of reference), and State vs. Process (e.g. in progressive aspect) underlie the mastery of certain IL subsystems.
- e. A certain Universal Accessibility Hierarchy may determine the interpretation of complex structures by IL users, as in the case of the *ask, tell, promise*.
- f. The influence of French is greater than that of the Bantu languages on the Zairean Learners' IL relative to English.
- g. Social and Psychological distance may affect both the rate and success in SLA, but not its natural route of development.

## CHAPTER 5

### RESULTS OF THE USE OF REFERENCE IN SLA

#### 5.1. Group Trends

##### 5.1.1. Group Range Variability

The first aspect of the learner's realisation of the article system to be investigated is the group range variability within the data (Andersen, 1978).

Four different types of articles are being examined:

1. The definite article *the*, i.e. when the referent is linguistically identifiable.
2. The definite article *the*<sub>2</sub>, i.e. when the referent is not linguistically identifiable;
3. The zero article *0*, i.e. in relation to the generic use and especially mass-count distinction;
4. The indefinite article *a, an*, i.e. in the case of first mention of a noun phrase in discourse.

Using a modified version of Andersen's (1978) Group Range Method, six group ranges are established to obtain an indication of the way in which the subjects' performance varies within each level of training in English. Table 5.1 gives the results of the group range method for the definite article *the*. These ranges were calculated using frequencies of correct responses on a number of items related to the function of *the* in English. These results indicate that most respondents are found in the fourth range (33.7 %) and fifth (29.5 %), totalling 63.2 % of correct uses. In other words, two thirds of the learners' performance are about average to above average. The other third (36.2 %) are still using these forms below average (i.e. first, second and third ranges). When we consider the 80% criterion, however, we realize that the university students are doing better with 47.6 % of the respondents ranging between 80 and 100%

accuracy rate. Then, follow the teacher trainees (30.3 %) and the high school students (24.2 %). Overall, 30% of the respondents (all levels combined) reached the > 80% criterion.

Table 5.1

Group Range Variability of Definite Article *THE*

Range of % correct	Level1 High school	Level2 Undergrad.	Level3 Teacher Tr.	Total
0 - 20 %	1 ( 1.5%)	2 ( 9.5%)	4 ( 5.3%)	7 ( 4.3%)
21 - 40 %	6 ( 9.1%)	2 ( 9.5%)	15 (19.7%)	23 (14.1%)
41 - 60 %	17 (25.8%)	3 (14.3%)	9 (11.8%)	29 (17.8%)
61 - 80 %	26 (39.4%)	4 (19.1%)	25 (32.9%)	55 (33.7%)
81 - 99 %	15 (22.7%)	10 (47.6%)	23 (30.3%)	48 (29.5%)
100%	1 ( 1.5%)	0 ( 0.0%)	0 ( 0.0%)	1 ( 0.6%)
TOTAL	66	21	76	163
Number > 80%	16 (24.2%)	10 (47.6%)	23 (30.3%)	49 (30.0%)

As regards the definite article *the*<sub>2</sub> (i.e. when the referent is not linguistically present) we find that the overall performance is much lower, at all levels. Looking at Table 5.2, we can see that more than half the respondents are found in the low 41-60% range (i.e. 53.4 %). What is even more puzzling is the fact that within this 41-60% range, the university students, who have been performing better on *the* earlier on, are doing worse than any other group. 61.9

% of the undergrads are scoring between 41 and 60% correct, compared with 56.6 % of the teacher trainees, and 47 % of the high school pupils. This group trend is further confirmed by the fact that none of the undergrads has reached the > 80% criterion as opposed to 7.9 % for the teacher trainees and 7.5 % for the high school pupils.

Table 5.2

Group Range Variability of Definite Article THE<sub>2</sub>

Range of % correct	Level1 High school	Level2 Undergrad.	Level3 Teacher Tr.	Total
0 - 20 %	0 ( 0.0%)	0 ( 0.0%)	1 ( 1.3%)	1 ( 0.6%)
21 - 40 %	6 ( 9.1%)	2 ( 9.5%)	8 (10.5%)	16 ( 9.8%)
41 - 60 %	31 (47.0%)	13 (61.9%)	43 (56.6%)	87 (53.4%)
61 - 80 %	24 (36.4%)	6 (28.6%)	18 (23.7%)	48 (29.5%)
81 - 99 %	5 ( 7.5%)	0 ( 0.0%)	6 ( 7.9%)	11 ( 6.7%)
100%	0 ( 0.0%)	0 ( 0.0%)	0 ( 0.0%)	0 ( 0.0%)
TOTAL	66	21	76	163
Number > 80%	5 ( 7.5%)	0 ( 0.0%)	6 ( 7.9%)	11 ( 6.7%)

As far as the *0* article is concerned, we find that the omission of the article is done correctly at a moderate rate. As shown in Table 5.3, the 61-80% range is the most dominant, at 42.3 % (all levels combined). Within this range, the teacher trainees are leading with 52.6 % of the correct responses, then both

the undergrads and the pupils are neck and neck (i.e. 33.3 %). Moreover, when we look at the > 80% criterion, we realize that this time the high school pupils are on top with 21.2 % of the correct use within the top range of 81-100%. The university students come second, with 19 %, then the teacher trainees with 14.5 %. Overall, the students' performance was rather modest with only 17.8 % reaching the > 80% criterion when all levels are combined.

Table 5.3

## Group Range Variability of Article ZERO

Range of % correct	Level1 High school	Level2 Undergrad.	Level3 Teacher Tr.	Total
0 - 20 %	11 (16.7%)	4 (19.0%)	3 ( 3.9%)	18 (11.0%)
21 - 40 %	12 (18.2%)	3 (14.3%)	4 ( 5.3%)	19 (11.7%)
41 - 60 %	7 (10.6%)	3 (14.3%)	18 (23.7%)	28 (17.2%)
61 - 80 %	22 (33.3%)	7 (33.4%)	40 (52.6%)	69 (42.3%)
81 - 99 %	14 (21.2%)	4 (19.0%)	10 (13.2%)	28 (17.2%)
100%	0 ( 0.0%)	0 ( 0.0%)	1 ( 1.3%)	1 ( 0.6%)
TOTAL	66	21	76	163
Number > 80%	14 (21.2%)	4 (19.0%)	11 (14.5%)	29 (17.8%)

The fourth article type is the indefinite article *a*, *an*. Table 5.4 gives these results, which suggest that once more the overall performance remains moderate. As with the *o* article, 42.3 % of the correct responses are found

within the 61-80 % range, the most dominant one as regards *a*, *an*. The university students are doing better (i.e. 52.4 %) than the other two groups. The teacher trainees are second (i.e. 44.7 %), and the pupils come third (i.e. 36.4 %). When we look at the > 80% criterion, we notice that this ordering is being slightly altered with the pupils ranking ahead (19.7 %), then the undergrads (19%), and finally the teacher trainees (15.8 %).

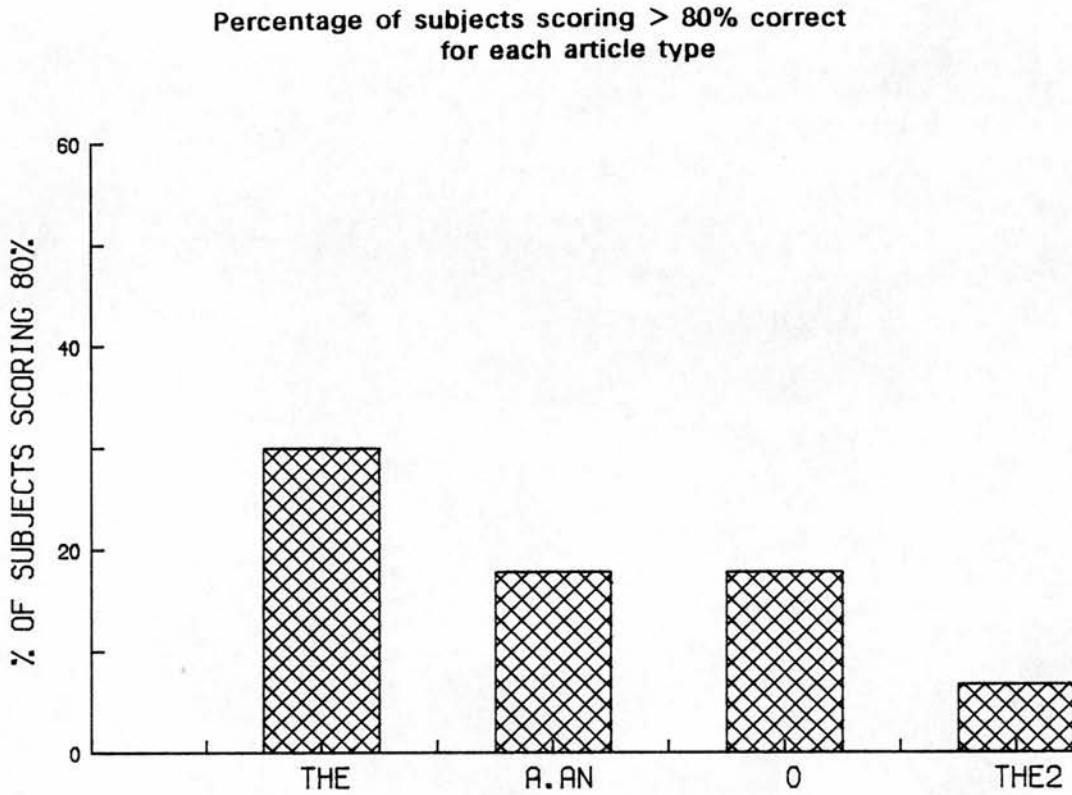
Table 5.4

Group Range Variability of the Indefinite Article *AN*

Range of % correct	Level1 High school	Level2 Undergrad.	Level3 Teacher Tr.	Total
0 - 20 %	0 ( 0.0%)	0 ( 0.0%)	1 ( 1.3%)	1 ( 0.6%)
21 - 40 %	11 (16.7%)	0 ( 0.0%)	10 (13.2%)	21 (12.9%)
41 - 60 %	18 (27.3%)	6 (28.6%)	19 (25.0%)	43 (26.4%)
61 - 80 %	24 (36.4%)	11 (52.4%)	34 (44.7%)	69 (42.3%)
81 - 99 %	13 (19.7%)	4 (19.0%)	11 (14.5%)	28 (17.2%)
100%	0 ( 0.0%)	0 ( 0.0%)	1 ( 1.3%)	1 ( 0.6%)
TOTAL	66	21	76	163
Number > 30%	13 (19.7%)	4 (19.0%)	12 (15.8%)	29 (17.8%)

Figure 5.1 gives the percentage of subjects scoring over 80% correct on each article type viz. *the* = 30%, *a*, *an* = 17.8%, *o* = 17.8%, and *the*<sub>2</sub> = 6.7%.

Figure 5.1



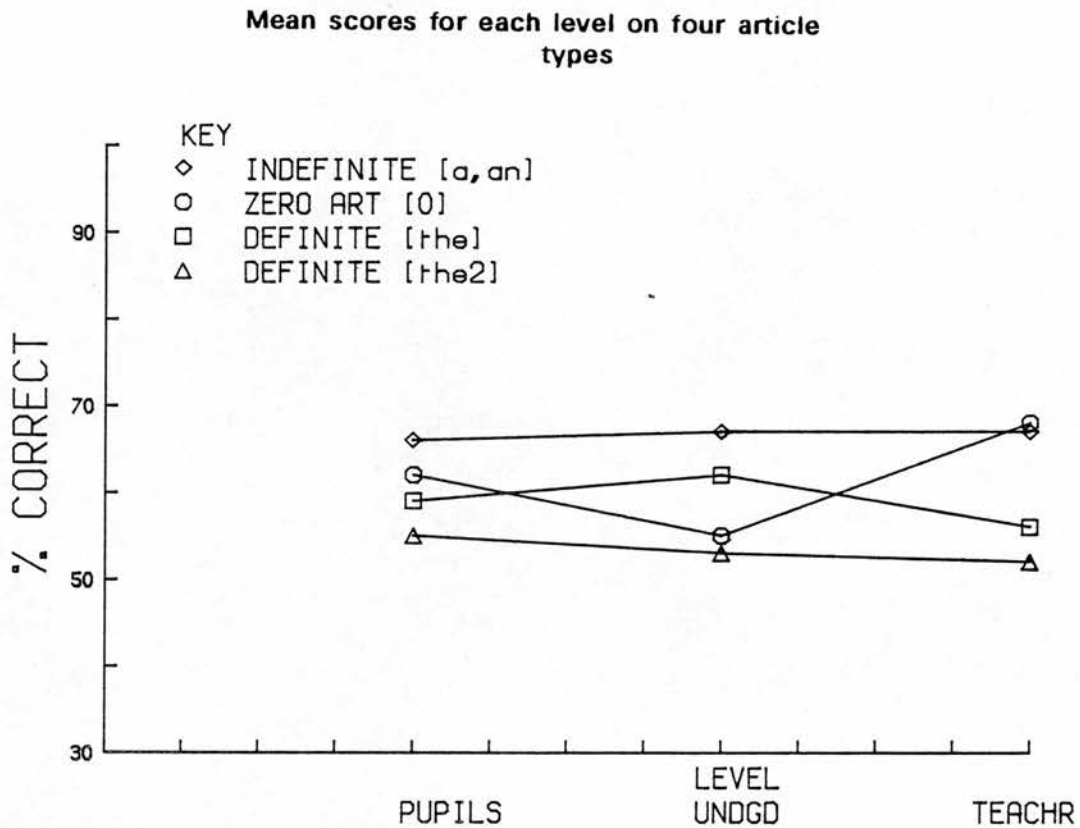
These results suggest that the definite article *the* is being learned faster than the indefinite article. However, in order to obtain a more accurate assessment of the respondents' performance we must find ways of dealing with the whole amount of variability within the data (not just frequencies). In the next subsection, we are going to calculate and compare mean scores to deal with this aspect of group behaviour.

### 5.1.2. Analysis of Variance and Scheffe Tests

In this section we report the results of the analysis of variance (ANOVA) conducted for comparing between the various means for the three levels of training in English and for the different article types. First we calculate the mean scores for each level, viz.  $\bar{X} = 61.09\%$  for level 1; 2)  $\bar{X} = 59.17\%$  for level 2; and 3)  $\bar{X} = 61.64\%$  for level 3. We then calculate the mean scores for each

of the four article types, viz. 1)  $\bar{X} = 58.5\%$  for *the* (i.e. definite article with identifiable referent); 2)  $\bar{X} = 54.2\%$  for *the*<sub>2</sub> (i.e. definite article with referent non-linguistically identifiable); 3)  $\bar{X} = 64.6\%$  for the zero article (i.e. generic use); and 4)  $\bar{X} = 67\%$  for the indefinite article *a, an*. Figure 5.2 gives the mean scores for each of the 3 levels on four article types.

Figure 5.2



These results indicate that all three groups (or levels) performed better on the indefinite article *a, an* than on either the zero article *0*, or the definite articles *the* and *the*<sub>2</sub>. The most difficult of all appears to be the definite article *the*<sub>2</sub>, followed by the other definite article *the* and then the *0* article. Conversely, the easiest of all is the indefinite article *a, an*. Table 5.5 gives the overall performance on the modified Cloze test task per article type for all levels combined.



Table 5.5

## Overall Performance per article type: All Levels combined

Article type	Mean	Standard Deviation	Number of observations
The	58.5	27.6	163
The <sub>2</sub>	54.2	26.6	163
Zero	64.6	37.2	163
A, An	67.0	28.3	163

The second step in our investigation of the performance on the article system is to examine whether the observed differences between levels' means are big enough to warrant our attention (i.e. to reach statistical significance). Also, we want to find out whether the degree of difficulty caused by the four articles are statistically significant. In order to conduct these analyses, we have formulated the following hypotheses:

1. There is no significant difference between the groups' means in their realisation of the articles.
2. There is no significant difference regarding the degree of difficulty caused by the four article types.
3. There is no significant difference between the linguistic environments such that the learners' performance could be affected by these differences.
4. There is no significant interaction of the combination of the three variables above, viz level, article type and linguistic environment.

These null hypotheses are to be tested in comparison with the following alternative hypotheses. First, those learners at level 3 (i.e. teacher trainees) are obviously expected to do better than both level 2 (i.e. undergrads) and level 1 (i.e. high school pupils), in view of the fact that the teacher trainees had

optimal exposure to English and could be expected to be highly motivated to learn and communicate in English. The university science students have had very restricted exposure and could be under pressure to concentrate first on their content subjects, thus, devoting less effort to improve their English. As regards the high school pupils, they would be expected to do worse than the teacher trainees. But the pupils might outperform the university students in spite of their (pupils') shorter length of exposure in number of years of training. One reason for this is that pupils in their final or penultimate year of high school are generally keen to learn as much as possible about any subject matter which is included in the State exams conducted in order to award the end of school 'Diplome d'Etat'. Another reason is that they do have more teaching hours per week than the university students do (i.e. five hours per week compared with one hour at the undergrad level).

Secondly, we expect the *0* article to be the most difficult in view of the problems facing French-speaking learners of English as regards the generic and mass-count distinction, as in e.g. *Milk and eggs are nutritious, but I prefer milk*. In French such a distinction is not grammatically marked and learners tend to supply *the* in these NPs contexts, where a *0* is required in Standard English. Thus, French-speaking (and other) learners of English often produce the incorrect utterance, viz. \**The milk and eggs are nutritious, but I prefer the milk*. (Le lait et les oeufs sont nourrissants, mais je prefere le lait). On the other hand, in spite of the conflicting findings in the literature (Borland 1984), the indefinite article is expected to be easier than definite articles, especially *the*<sub>2</sub>, the referent of which is not linguistically present in the text. The indefinite article operates in the same way in both English and French to signal the first mention of a referent, a topic, or NP in discourse; and, thus, *a*, *an* would not cause difficulty.

Thirdly, we expect the subject environment to be easier than the object environment. According to Givon (1984) and Huebner (1983), the NP in the subject or agent position is normally the 'topic', which in turn represents given information. On the other hand, the NP in the object position is usually, the 'comment', i.e. new information which is not yet established, and thus may be difficult to identify correctly.

Fourth, we expect some interaction between these three variables; viz. level, article and linguistic environment, although we want to know how exactly such an interaction would be. According to some studies in the literature (Meisel, Clahsen and Pienemann 1981), the linguistic development of L<sub>2</sub> learners cannot be defined in linear terms, i.e. SLA is a multi-dimensional process, implying that learner strategies may change depending on the level of proficiency or certain subsystems of his IL may be differently internalized by the learner. The prepositional phrase (PP) environment might be located somewhere between the subject and object environments. The analysis of variance (ANOVA) was performed using the Biomedical Computer Program Statistical Software (1983). The actual procedure utilized to test the above null hypotheses was the BMDP2V 'ANOVA and Covariance with Repeated Measures.'

Our design involved 1) English Training (LEVENG) as the independent variable with three levels (i.e. level 1 = school, level 2 = undergrad, level 3 = postgrad or teacher trainees); 2) Article (ART) as the dependent variable with four types (as repeated measures) and three environments (as nested within each article type); thus the dependent variable ART consisted of  $4 \times 3 = 12$  variables.

Table 5.6

## ANOVA Results for 4 articles in 3 Environments by levels

SOURCE Main effects:	SS	d.f.	M.S.	F	P
Level of T.	1199.887	2	599.943	0.30	0.738
Articles	31531.419	3	10510.473	7.85	0.000**
Environments	28536.494	2	14268.247	20.60	0.000**
2-Way Inter (AE)	134420.301	6	22403.383	34.75	0.000**
3-Way Inter. (AEL)	12678.036	12	1056.503	1.64	0.07
** p < .01					

The ANOVA results are given in Table 5.6. First of all, the main effect for 'level of training' in English shows a non-significant F-ratio value (  $F = 0.30$ ), suggesting that there is no significant difference between the groups' means in their performance on the articles as a whole; thus, this hypothesis will not be pursued any further. In addition, the second main effect for articles reveals a highly significant F-value ( $F = 7.85$ ;  $df = 3, 480$ ;  $p < .0001$ ). These results indicate that there are considerable differences between the articles, and thus, we are able to reject the null hypothesis of equality among them. In order to test the differences between pairs of articles, the Scheffe procedure was conducted for 'post hoc comparisons' (Winer, 1971: 185; Hatch and Farhady, 1982: 140). The Scheffe tests results for comparison between pairs of articles are given in Table 5.7. The definite articles *the* / *the*<sub>2</sub> are significantly different from the indefinite article *a*, *an* ( $p < .01$ ). In other words, the respondents found the indefinite article easier than either *the* or *the*<sub>2</sub>. The *0* article differs significantly from *the*<sub>2</sub> only, but not with any other type. This was a somewhat

mixed finding suggesting that contrary to our second alternative hypothesis, the *0* article or generics would be more difficult than *the* definite article. In fact, the *0* article appears to be much easier than *the*<sub>2</sub>, but of equal difficulty with *the*. Moreover, the comparison between *the* and *the*<sub>2</sub> does not reach any significance.

Table 5.7

## Scheffe Tests for Comparison Between Pairs of Articles

mean=	THE <sub>2</sub>	THE	O	AN, A
mean= 26503.8	26503.8	28606.5	31586.4	32763.0
THE <sub>2</sub>	-		5085.6**	6259.2**
28606.5		-		4156.5**
THE				
** p<.01      t-crit.= 3875.9				

Regarding the third main effect, i.e. linguistic environments, the ANOVA results are highly significant ( $F = 20.6$ ;  $df 2, 320$ ;  $p < .0001$ ). This was a very encouraging finding, leading to the rejection of our third null hypothesis. In this case, the linguistic environments did affect the learners' performance on the use of reference significantly. Appropriate Scheffe tests (see Table 5.8) show that the prepositional phrase is the easiest; then follows the subject / agent NP environment, and the object NP environment is the most difficult of all. These results are highly significant ( $p < .01$ ) and support our third alternative hypothesis i.e. the comment position is the most likely to cause difficulty. Presumably the PP degree of difficulty is due to the fact that the PP constituent is at the interface of syntax and semantics; and, thus, the learners may be puzzled at times just as their task might be made easier at others.

Table 5.8

## Scheffe Tests for Comparison Between Pairs of Environments

	OBJ+NP	AGT+NP	PREP.P.
mean=	35338.4	41728.0	42314.8
mean= 35338.4 (OBJ+NP)	-	6389.6 **	6976.4 **
** p<.01      t-crit.= 2882.62			

Next, we examined whether there was an interaction effect between the three main effects. Table 5.9 gives the results in relation to the possible interaction effects between levels of training, articles and environments. No significant interaction effect was found between article types and levels of training ( $F = 1.46$ ;  $df = 6, 480$ ; n.s.). This suggests that the articles are not just difficult within themselves, they are also difficult between all subjects, irrespective of their level. In addition, there was no interaction effect between environments and levels. However, there was a highly significant 2-way interaction effect between articles and environments ( $F = 34.75$ ;  $df = 6, 960$ ;  $p < .0001$ ). These results indicate that some articles are indeed more difficult to supply in certain linguistic environments than others. In other words, certain combinations of 'article with environment' do affect the performance on the use of reference more seriously than others, for all subjects irrespective of their level. This suggestion is borne out by the lack of significance of a 3-way interaction effect related to levels, articles and environments ( $F = 1.64$ ;  $df = 12, 960$ ;  $p < .07$ ). Since there is a significant interaction effect between articles and environments, we have to examine which combination(s) of 'article with environment' was causing greater difficulty.

Table 5.9

Scheffe Tests for Articles with Environments

	THE2	O	THE	AN	THE	THE	THE2	AN	THE2	O	O	AN
	+	+	+	+	+	+	+	+	+	+	+	+
	OBJ	AGT	OBJ	PREP.P	PREP.P	AGT	AGT	OBJ	PREP.P.	OBJ	PREP.P.	AGT
mean =	5797.91	7799.55	8658.56	8839.49	9848.46	10115.78	10272.89	10355.39	10440.15	10529.80	13248.64	13599.00
(THE2 + OBJ)												
5797.91			2860.65	3041.58	4060.55	4317.87	4475.98	4557.48	4642.24	4731.89	7450.70	7801.00
(THE2 + OBJ)			**	**	**	**	**	**	**	**	**	**
7799.55					2048.91	2316.23	2474.34	2555.84	2640.60	2730.25	5449.09	5799.45
(O + AGT)					*	**	**	**	**	**	**	**
8658.56											4590.08	4940.44
(THE + OBJ)											**	**
8839.49											4409.15	4759.51
(AN + PREP.P.)											**	**
9848.46											3400.18	3750.54
(THE + PREP.P.)											**	**
10115.78											3132.86	3483.22
(THE + AGT)											**	**
10272.89											2974.75	3325.11
(THE2 + AGT)											**	**
10355.39											2893.25	3243.61
(AN + OBJ)											**	**
10440.15											2808.49	3158.85
(THE2 + PREP.P.)											**	**
10529.80											2718.84	3069.20
(O + OBJ)											**	**

\*\* p < .01 t-crit. = 2282.19  
 \* p < .05 t-crit. = 2034.28

Appropriate Scheffe tests demonstrate that the most difficult type of article is the definite *the*<sub>2</sub> (i.e. when the referent is not linguistically identifiable), and more specifically so when the *the*<sub>2</sub> occurs before an object NP. The easiest types are the indefinite article *a, an* in the subject NP environment and the zero *0* article before a PP. Why these combinations should be more difficult and others easier is a matter of great interest for SLA researchers. At this stage one can only speculate that this may be due to:

1. Language specific properties of the target language (i.e. linguistic markedness).
2. Certain predictions underlying the Wave model mechanism of linguistic change, advocated by such researchers as Bailey (1973), Dickerson (1976), and Zobl (1983, 1984).
3. Universals of 'discourse-pragmatics or pre-syntactic communication' of the type proposed by Givon (1979, 1984); or indeed,
4. Random variability in the data - although this seems totally implausible in view of the sensitiveness of the Scheffe procedure (Winer, 1971).

## 5.2. Individual Trends

### 5.2.1. Scalogram Analysis for Articles

In the previous section, we have looked at the overall differences between groups' means on the learner's use of definite and indefinite reference. Both ANOVA and Scheffe tests have enabled us to find out how much the subjects have learned in this area. Our main concern in this section is somewhat different. What we want now is to be able to say that this learner or this group of learners are using this article at stage 1; this is the article they tend to use at stage 2; and that is the article they supply at stage 3, etc. The Scalogram analysis was conducted with the subprogramme known as the Guttman scale (Nie et al. 1975), using the cut-off point of > 80% criterion. The four article



types were ordered following the pattern provided by Scheffe tests. As we have just said (in the last section), the easiest article is the indefinite *a, an*; followed by the zero; then, the definite *the*; and finally, the most difficult of all viz. the definite *the*<sub>2</sub> (i.e. with no linguistically identifiable referent). Thus, the pattern predicted by Scheffe is as follows:

$$a, an > 0 > the > the_2$$

However, when the articles were analysed by Guttman scale in this fixed order, the scale failed to reach significance (coefficient of reproducibility = .72; coefficient of scalability = n.s.). Upon closer examination of the correlation coefficient supplied by Guttman Scale Statistics, we found that the 0 article did not fit into this set (Yule's Q = -0.13 for zero article). For this reason, the 0 article was removed from the scale; thus, only the remaining three items are included in subsequent analyses by Guttman scale. The results in relation to these three articles are given in Table 5.10 (coef. rep. = 0.87; coef. scal. = 0.39). Although the scale itself is not statistically significant (i.e. coef. of scal. must be at least 0.60), there appears to be a more positive interrelationship between these 3 articles, as indicated by the coef. of rep. = 0.9 (i.e. if we used one decimal place) and Yule's Q correlations (not shown in the table).

Table 5.10

## Guttman Scale for 3 Articles

ITEM	THE2		A, AN		THE		TOTAL
RESPONSE	0	1	0	1	0	1	
SCALE	ERR		ERR		ERR		
3	0	5	0	5	0	5	5
2	13	12	9	16	3	22	25
1	31	6	24	13	19	18	37
0	96	0	96	0	96	0	96
SUMS	140	23	129	34	118	45	163
PCTS	86	14	79	21	72	28	
ERRORS	0	18	9	13	22	0	62

163 cases processed  
 0 (or 0.0 PCT) were missing

## STATISTICS

1. Coefficient of Reproducibility = 0.87
2. Minimum Marginal Reproducibility = 0.79
3. Percent Improvement = 0.08
4. Coefficient of Scalability = 0.39

The Guttman scale results also predict an order of difficulty slightly different from that of Scheffe tests. However, this is quite conceivable since the Guttman procedure takes into account only those subjects who have reached the > 80% criterion, whereas Scheffe (or ANOVA) deals with the total amount of variability within the data. According to the Guttman scale results, the preferred order is as follows:

*the* > *a, an* > *the*<sub>2</sub>

. As we saw while dealing with group range variability (section 5.1.1), the definite article *the* alone accounted for 30% of the subjects within the top range (i.e 80-100%), whereas only 17.8% were found in this range as regards performance on the indefinite article *a, an*. The definite article *the*<sub>2</sub> accounted for 6.7%. If the article *the*<sub>2</sub> was removed or incorporated within the article *the* (since they did not seem to differ substantially in quantitative terms), then the overall order of difficulty would be *the* > *a, an*. In other words, the definite article would appear to have been learned at a faster rate by these learners. This finding corresponds to what Borland (1984) and Okanlawon (1984) suggest while dealing with four groups of immigrant pupils learning English in Australia and Nigerian school pupils respectively.

The proportion of subjects at each level who had reached the > 80% criterion for each article is summarised in Table 5.11.

**Table 5.11**

**Proportion of Subjects Reaching the 80% Criterion**

Levels	Article type:		
	THE	A, AN	THE2
1. Pupils	16/66 (24.2%)	13/66 (19.7%)	5/66 ( 7.5%)
2. Undergraduates	10/21 (47.6%)	4/21 (19.0%)	0/21 ( 0.0%)
3. Teacher trainees	23/76 (30.0%)	12/76 (15.8%)	6/76 ( 7.9%)
All levels	49/163 (30.0%)	29/163 (17.8%)	11/163 ( 6.7%)

In addition, we constructed (manually) a number of Bimodal Implicational Scales (Hatch and Farhady, 1982) in order to find out which were the most favoured orderings per level. These results are reported in Tables 5.12, 5.13, 5.14, and 5.15.

Table 5.12

**Implicational Scale for Second Year Teacher Trainee  
Students (n=43)**

S.No.	THE	A, AN	THE2	O (excluded)
7	+	+	+	-
6	+	+	+	-
41	+	+	-	-
42	-	+	-	-
24	-	+	-	-
2	-	+	-	-
15	+	-	+	-
5	+	-	+	-
29	+	-	-	-
39	+	-	-	+
35	+	-	-	+
38	+	-	-	+
26	-	-	+	-
43	-	-	-	+
23	-	-	-	-
37	-	-	-	-
40	-	-	-	-
10	-	-	-	+
18	-	-	-	+
8	-	-	-	-
31	-	-	-	-
32	-	-	-	-
20	-	-	-	-
33	-	-	-	-
27	-	-	-	-
17	-	-	-	-
13	-	-	-	+
9	-	-	-	-
21	-	-	-	-
19	-	-	-	+
1	-	-	-	-
11	-	-	-	-
3	-	-	-	+
28	-	-	-	-
30	-	-	-	-
36	-	-	-	-
4	-	-	-	+
25	-	-	-	+
34	-	-	-	-
14	-	-	-	-
22	-	-	-	-
12	-	-	-	+
16	-	-	-	+
Correct	9	6	5	13 = 20
Errors	4	2	3	= 9

Coefficient Rep. = 0. 93  
Coefficient Scal. = 0. 92

As shown in Table 5.12, it is clear that *0* article did not fit into the scale and was removed from further analyses. This decision was reached upon inspection of this trial implicational scale conducted for only 43 subjects, viz. the most advanced of all (i.e. those who were in their second year of the teacher's college). Table 5.12 shows that the subjects who score pluses (+) for *0* article are almost exclusively the ones who score minuses (-) for *the*, *a*, *an*, and *the*<sub>2</sub>. These include subjects no 43 to 16 in the table; and represent 77% of all the 'correct' uses of the zero article at the > 80% criterion (i.e. 10 out of 13 pluses). Thus, the *0* is not included in the calculation of the statistical significance of this implicational scale, which was found to be high (coef. rep. = 0.93; coef. scal. = 0.92).

These results suggest that when the subjects left a blank on their answer sheet this meant that either they interpreted the intended meaning of *0* article correctly; or they were not so sure about it; or, indeed, they did not know the correct answer at all. Because of this ambiguity inherent in the use of generics, and the *0* article in particular, a more detailed investigation of this structure was considered unwarranted, and the study of *0* was not pursued further. Presumably intuitional data might be a more appropriate way of investigating the form-function relationship involved in the use of generic reference.

Table 5.13

Implicational Scale for the Pupils' Use of Reference  
(n=66)

S.No.	THE	A, AN	THE2
157	+	+	+
160	+	+	-
141	+	+	-
116	+	+	-
135	+	+	-
155	-	+	+
142	-	+	-
130	-	+	-
106	-	+	-
149	-	+	-
136	-	+	-
119	-	+	-
158	-	+	-
162	+	-	+
165	+	-	+
148	+	-	+
147	+	-	+
144	+	-	+
112	+	-	+
137	+	-	-
109	+	-	-
131	+	-	-
114	+	-	-
145	+	-	-
164	-	-	+
127	-	-	-
167	-	-	-
151	-	-	-
138	-	-	-
163	-	-	-
108	-	-	-
118	-	-	-
103	-	-	-
110	-	-	-
161	-	-	-
134	-	-	-
115	-	-	-
139	-	-	-
Correct	16	13	9 = 38
Errors	9	0	8 = 17

Coefficient rep. = 0.91

Coefficient scal. = 0.89

Table 5.14

**Implicational Scale for the Undergraduates' Use of  
Reference (n=21)**

S.No.	THE	A, AN	THE2
96	+	+	+
84	+	+	-
95	+	-	+
77	-	+	-
80	-	+	-
89	+	-	-
85	+	-	-
79	+	-	-
86	+	-	-
94	+	-	-
91	+	-	-
87	+	-	-
97	-	-	-
88	-	-	-
83	-	-	-
90	-	-	-
81	-	-	-
92	-	-	-
78	-	-	-
82	-	-	-
93	-	-	-
Correct	10	4	2 = 16
Errors	2	1	1 = 4

Coefficient rep. = 0.90

Coefficient scal. = 0.87



Table 5.15

**Implicational Scale for the Teachers' Use of  
Reference (n=76)**

S.No.	THE	A, AN	THE2
7	+	+	+
6	+	+	+
74	+	+	-
41	+	+	-
46	+	+	-
48	+	+	-
54	+	+	-
5	+	-	+
15	+	-	+
47	+	-	+
67	+	-	+
71	+	-	+
29	+	-	-
39	+	-	-
65	+	-	-
66	+	-	-
68	+	-	-
55	+	-	-
72	+	-	-
35	+	-	-
38	+	-	-
49	+	-	-
69	+	-	-
42	-	+	-
24	-	+	-
2	-	+	-
64	-	+	-
61	-	+	-
26	-	-	+
44	-	-	-
53	-	-	-
58	-	-	-
62	-	-	-
75	-	-	-
43	-	-	-
23	-	-	-
70	-	-	-
76	-	-	-
Correct	23	12	8 = 43
Errors	6	10	6 = 22

Coefficient rep. = 0. 90

Coefficient scal = 0. 88

In addition, Tables 5.13, 5.14, and 5.15, give the most favoured 'acquisitional' orders as regards the remaining three articles, by each level of training. The criterion for accuracy was > 80% and the statistical results were all significant, as follows:

Table 5.16

Most Favoured Orders of Accuracy for Article Types  
by Levels.

Level	Implicational pattern	Significance level
1. Pupils	THE > A, AN > THE2	coef rep = .91 coef scal = .89
2. Undergrad.	THE > A, AN > THE2	coef rep = .90 coef scal = .87
3. Teacher train	THE > A, AN > THE2	coef rep = .90 coef scal = .88

### 5.2.2. Scalogram Analysis for Articles within 3 Linguistic Environments

In the previous section, it was suggested that the definite article *the* was being learned faster, followed by indefinite *a*, *an*, and then the definite *the*<sub>2</sub> (at least as far as the 80% criterion was concerned). In this section, Scalogram analysis is aimed at discovering how the subjects' performance on each article could be characterized according to the position of the Noun Phrase. Various implicational scales have been constructed indicating the most favoured orderings of the different articles within environments per level. These results are given in Tables A - I (see appendix). The three article types were examined separately, with particular reference to 1.) environment, and 2) levels. First of all, as regards the definite article *the*, the prepositional phrase (PP) environment appeared to be the most favoured across all three levels. (Tables A, B, C, and D,

in appendix); then followed the subject agent (AGT) position; and, the least favoured was the object (OBJ) position. These results were all statistically significant and summarised as follows:

Table 5.17

**Implicational Patterns for the use of Definite Reference *THE*  
at three levels**

Level	Implicational pattern	Significance level
1. Pupils	THE+PP>THE+AGT>THE+OBJ	coef rep = .91 coef scal = .86
2. Undergr.	THE+PP>THE+AGT>THE+OBJ	coef rep = .90 coef scal = .82
3. Teacher tr.	THE+PP>THE+AGT>THE+OBJ	coef rep = .92 coef scal = .87

Secondly, in relation to the indefinite article *a, an* the most favoured accuracy order for the three linguistic environments was the same for all three levels of training in English. These results are given in Tables D to F (see appendix)

Table 5.18

**Implicational Patterns for the Use of Indefinite Reference  
at three levels**

Level	Implicational pattern	Significance level
1. Pupils	AN+AGT > AN+OBJ > AN + PP	coef rep = .93 coef scal = .87
2. Undergrad.	AN+AGT > AN+OBJ > AN + PP	coef rep = .90 coef scal = .80
3. Teacher train.	AN+AGT > AN+OBJ > AN + PP	coef rep = .90 coef scal = .78

Thirdly, regarding the definite article *the*<sub>2</sub> the predicted order of difficulty for

the three environments was the same for both the teacher- trainees and school pupils, but this order was different for the undergraduate science students. These results are given in Tables G, H, and I (see appendix), and were all highly significant. Table 5.19 gives a summary of the most favoured orderings per level in relation to the definite article *the*<sub>2</sub>.

**Table 5.19**

**Implicational Patterns for the Use of Definite  
Reference THE<sub>2</sub>  
at three levels**

Level	Implicational pattern	Significance level
1. Pupils	THE2+AGT>THE2+PP>THE2+OBJ	coef rep = .91 coef scal = .88
2. Undergrad.	THE2+PP >THE2+AGT>THE2+OBJ	coef rep = .87 coef scal = .84
3. Teacher tran.	THE2+AGT>THE2+PP>THE2+OBJ	coef rep = .94 coef scal = .93

From these results, it appeared that, as far as the article system was concerned, there was a variable implicational difficulty order according to which in some cases (or at certain levels) the articles were easier in the agent / subject NP position or PP position. However, the articles under investigation were almost invariably more difficult in the OBJ NP position as compared with the subject NP position. Thus, the following pattern for best order of difficulty for the three environments was established:

$$\text{AGT NP} \geq \text{PP NP} \geq \text{OBJ NP}$$

### 5.3. Learner Language Behaviour

In the previous sections we reported on the groups' and individuals' correct suppliance of the articles. Using both ANOVA / Scheffe tests and Scalogram analysis, we established the variability and systematicity related to the subjects' use of reference in obligatory contexts. However, in order to obtain a more complete picture of the learner's internalized IL system, we have to take into account not only the correct suppliance of the articles but their incorrect uses as well. In this section we will first report on the overall suppliance of *the*, *the*<sub>2</sub> and *a*, *an* for all levels of training. Secondly, we will examine the realization of the article system (or use of reference) in grammatical contexts, i.e. incorrect uses of *the*, *the*<sub>2</sub>, and *a*, *an* for all groups. Since one of the incorrect uses of any of these articles would involve its omission, the *0* article will also be included as necessary. Thirdly, we will look at the suppliance of correct or incorrect forms by each group, in three linguistic environments.

Table 5.20

Percent Correct Use of the Articles  
in Obligatory Contexts

		Correct article			
		THE <sub>2</sub>	THE	O	A, AN
Actual response					
	THE	54.2	58.5		
	O			64.6	
	A, AN				67.0

### 5.3.1. The Use of Definite and Indefinite Reference In Obligatory Contexts

Table 5.20 provides the overall mean scores pertaining to each of the article types, viz, *the*, *the*<sub>2</sub>, *a*, *an* and *0*. As shown in this table, the overall performance on the indefinite article *a*, *an* was the highest ( $\bar{X} = 67\%$ ), whereas the performance on the definite article *the*<sub>2</sub> was the lowest ( $\bar{X} = 54.2\%$ ). Meanwhile the zero article *0* came second highest ( $\bar{X} = 64.6\%$ ); and the definite article *the* was third ( $\bar{X} = 58.5\%$ ). We have already established the differences between these means by ANOVA, Scheffe tests (section 5.1.2) and we shall not elaborate on those analyses any further. The only point worth remembering in relation to *0* is that the omission of an article is rather ambiguous since it may mean a correct response or an erroneous deletion (i.e. acting as a distractor, in the Multiple Choice testing sense). In the next paragraph, we will examine the proportion of incorrect use that was due to this ambiguous phenomenon of deletion or to a deliberate choice of an erroneous article type.

### 5.3.2. The Use of the Definite and Indefinite Reference in Ungrammatical Contexts

Table 5.21 provides the results in relation to the incorrect uses of the different articles by the different groups combined. The correct uses are exactly those given earlier (see Table 5.20). The incorrect uses (Table 5.21) are referred to as 'actual responses'. These indicate the relative incidence of each alternative choice (i.e. *0* and *the* or *a*, *an*) in the ungrammatical use.

Table 5.21

Percent Correct Realization of the Article System  
in both Obligatory and Ungrammatical Contexts (all levels combined)

	Correct article			
	THE <sub>2</sub>	THE	O	A, AN
Actual response	54.2%	58.5%	64.6%	67.0%
THE	-	-	20.8	17.3
O	34.6	27.2	-	15.7
A, AN	11.2	14.3	14.6	-

These results suggest that, as regards *the*<sub>2</sub> article, the incorrect uses accounted for 45.8% of the overall performance (i.e. 34% for *o* and 11.2% for *a, an*). 76% of the incorrect (actual) responses were due to deletion rather than to the confusion between *the*<sub>2</sub> and *a, an*. Only 24% could be attributed to the confusion between *the*<sub>2</sub> and *a, an* leading to an incorrect response. In addition, since *a, an* occurred less than 33% in this environment, we argue that the 24% rate of occurrence might have been due to chance alone; and thus, we disregarded its importance.

Secondly, in connection with the *the* article, we find that 41.5% of the performance is ungrammatical (i.e. 27.2% for *o* and 14.3% for *a, an*). 66% of the incorrect/actual uses were due to the deletion of the article, whereas only 34% were due to the wrong choice of *a, an*. Again we feel that this replacement of *the* by *a, an* is close to chance occurrence; and so, we retained for deletion (i.e. *o*) as the relevant factor leading to ungrammatical use).

Thirdly, concerning the *a, an* article, we find that both the incorrect deletion of *a, an* and its replacement by *the* accounted for 33%, which is exactly what

one would have expected in the event of both deletion (*Ø*) and replacement (*the*) being resorted to by chance. So far, these results suggest that the only instances in which deletion was not resorted to by chance (i.e. more than 33%) were in relation to the articles *the*<sub>2</sub> (i.e. 76% of wrong deletion) and *the* (i.e. 66% of wrong deletion). In both cases however, the indefinite article did not interfere with the subjects' performance in a significant way. We would then argue that the *Ø* was the manifestation of the avoidance phenomenon which was probably due to the learners' restricted linguistic knowledge.

Fourthly, as far as the *Ø* article is concerned, Table 5.21 shows that, 35.4% of the time, one of the two articles (viz. *the* or *a, an*) was supplied instead of the generic *Ø* article.

These results suggest that when *the* or *a, an* was supplied in ungrammatical contexts, this was not due to sheer guessing. If this were the case, then all these instances of actual encroachment on the function of *Ø* by the use of the other article types would fall below the 33% limit for chance occurrences. Instead, we obtained 35.4% of occurrences of incorrect use. This figure gives us an indication of the likely role that both transfer and overgeneralization may have played in causing the erroneous use of *the* or *a, an* in contexts where the generic *Ø* was required in Standard English. In other words, there was some positive initiative (or meta-linguistic awareness) underlying the suppliance of *the* and *a, an*. Sometimes this initiative led to correct use; but, also, sometimes it landed the learners into trouble (i.e. they produced an incorrect response). In addition, the correct use of *the* instead of *Ø* might be due to the interference of French (as in e.g. *\*I like the milk*, meaning to say *I like milk* or *J'aime le lait* in French). However, the incorrect uses could also be due to generalization of *the* to the function of the generic article (viz. to mark definite or non-definite,



non-specific referents, assumed known to the hearer; in short, (+/-D - SR + HK). By comparison, the definite article is normally used to mark only definite, specific referents which are assumed known to the hearer (in short, (+D + SR + HK)).

As regards the incorrect use of *a, an* instead of *le, la*, overgeneralization is the most likely error cause. The reason for this is that the indefinite article functions in a similar way in both French and English, which rules out transfer. More specifically, the indefinite article is used to mark referents which are introduced into the discourse for the first time (i.e. non-definite, specific NPs which are not assumed known to the hearer; or in short (-D + SR - HK). Thus, the confusion or error may be due to both form and function of *a, an*. The form can be used for both indefinite and generic articles. Regarding the function, learners may consider certain referents as specific and assumed known to the teacher (or other intended audience), while these referents are not so specific in textual discourse. Thus, both language competence and shared knowledge may account for most of these incorrect uses.

Table 5.22

**Correct and Incorrect Uses of *THE* in 3  
Linguistic Environments**

Structure within environment	5th form High School	6th form High School	1-2 year Under-graduate	1st year Teacher College	2nd year Teacher College
- AGT NP					
correct	.56	.70	.60	.55	.71
incorrect					
O	.23	.16	.23	.33	.17
A, AN	.21	.14	.17	.12	.12
- OBJ NP					
correct	.50	.58	.57	.48	.58
incorrect					
O	.36	.24	.24	.40	.29
A, AN	.14	.18	.19	.12	.13
- PP					
correct	.58	.73	.63	.47	.69
incorrect					
O	.30	.20	.29	.45	.22
A, AN	.12	.07	.08	.08	.09
TOTAL					
correct	.55	.68	.60	.52	.67
incorrect					
O	.28	.19	.24	.37	.21
A, AN	.17	.13	.16	.11	.12

(Frequencies expressed in proportions, i.e. observed correct and incorrect responses relative to maximum number of expected responses within each environment.)

Table 5.23

**Correct and Incorrect Uses of THE<sub>2</sub> in 3  
Linguistic Environments.**

Structure within environment	5th form High School	6th form High School	1-2 year Under- graduate	1st year Teacher College	2nd year Teacher College
- AGT NP					
correct	.56	.74	.60	.55	.73
incorrect					
O	.32	.20	.23	.41	.24
A, AN	.12	.06	.17	.04	.03
- OBJ NP					
correct	.25	.48	.37	.44	.23
incorrect					
O	.53	.35	.54	.41	.69
A, AN	.22	.17	.09	.15	.08
- PP					
correct	.66	.70	.66	.56	.66
incorrect					
O	.26	.19	.23	.37	.27
A, AN	.08	.11	.11	.07	.07
TOTAL					
correct	.55	.66	.58	.53	.59
incorrect					
O	.33	.23	.30	.39	.35
A, AN	.12	.11	.12	.08	.06

(Frequencies expressed in proportions,  
i.e. observed correct and incorrect responses relative to maximum  
number of expected responses within each environment.)

Table 5.24

**Correct and Incorrect uses of *A, AN* in  
3 Linguistic Environments**

Structure within environment	5th form High School	6th form High School	1-2 year Under-graduate	1st year Teacher College	2nd year Teacher College
- AGT NP					
correct	.86	.83	.76	.79	.88
incorrect					
O	.00	.14	.14	.07	.06
THE	.14	.03	.10	.14	.06
- OBJ NP					
correct	.58	.72	.67	.58	.67
incorrect					
O	.34	.15	.11	.34	.25
THE	.08	.13	.22	.08	.08
- PP					
correct	.48	.50	.59	.57	.58
incorrect					
O	.16	.41	.26	.17	.10
THE	.36	.09	.15	.26	.32
TOTAL					
correct	.57	.65	.65	.58	.66
incorrect					
O	.26	.24	.16	.28	.18
THE	.17	.11	.19	.14	.16

(Frequencies expressed in proportions, i.e. observed correct and incorrect responses relative to maximum number of expected responses within each environment.)

### 5.3.3. The Use of Definite and Indefinite Reference in Three Linguistic Environments

We now turn to the investigation of the way in which the three article types were being used in three linguistic environments, viz. the subject/agent NP, the object NP and the Prepositional Phrase. Both correct and incorrect uses are being accounted for. Tables 5.22, 5.23 and 5.24 have given the proportions

in relation to the correct and incorrect uses of *the* and *the*<sub>2</sub> and *a*, *an* in three linguistic environments. In order to obtain a better assessment of the variability in the learning process of these forms, the subjects were subdivided into five groups (i.e. the 5th form, and 2) 6th form high school, 3) the university students, 4) the 1st year and 5) 2nd year of Teacher Training college). These groups represent, in effect, five different points in time in the learning process of the TL. From these results, it appears that the mastery of the NP reference system has been taking place, variably but systematically, within different environments at any one time. Similarly, when we examine the kind of change which has been taking place across the different levels or groups, we notice a highly variable pattern of behaviour as well. Apart from the indefinite article *a*, *an* in the subject NP environment (which has reached the  $\geq 80\%$  criterion), all the other articles within particular environments are still being realized variably (i.e. between 79% and 55%) or in a pre-systematic way (i.e. between 54% and 0%). In addition, in all environments, most incorrect uses are instances of deletion.

Table 5.25 summarizes the correct use of each article in three linguistic environments by groups. These results reveal variable implicational patterns such that, as expected, group 5, then 4, then 3 etc. was learning the articles faster than its nearest group from right to left. In addition, overall, the indefinite article was being mastered faster than the definite *the*. However, no implicational relationships have been found among the five groups as regards their performance on *the*<sub>2</sub>. Presumably this shortcoming could be circumvented by looking at individual's behaviour. These group trends suggest that this article sub-type (i.e. *the*<sub>2</sub>) was of equal difficulty for all. Concerning the different environments, we find that the subject NP position was the easiest, and the object NP the most difficult; the PP environment was rather unpredictable.

Table 5.25

**Correct use of each article within three linguistic environments by levels**

Structure within environment	LEVEL 1		LEVEL 2	LEVEL 3	
	5th form High School	6th form High School	1-2 year Under-graduate	1st year Teacher College	2nd year Teacher College
AN - AGT NP	.86 ⊕	.83 ⊕	.76 x	.79 x	.88 +
- OBJ NP	.58 x	.72 x	.67 x	.58 x	.67 x
- PP	.48 -	.50 -	.59 x	.57 x	.58 x
THE - AGT NP	.56 x	.70 x	.60 x	.55 x	.71 x
- OBJ NP	.50 -	.58 ⊗	.57 ⊗	.48 -	.58 x
- PP	.58 ⊗	.73 ⊗	.63 ⊗	.47 -	.69 x
THE2 - AGT NP	.56 x	.74 x	.60 x	.55 x	.73 x
- OBJ NP	.25 -	.48 -	.37 -	.44 -	.23 -
- PP	.66 ⊗	.70 ⊗	.66 ⊗	.56 ⊗	.66 ⊗

Key: + Categorical use of article (100 - 80%)  
 x Variable use of article (79 - 55%)  
 - Pre-systematic use (54% or less)  
 O circled marks = errors in scaling patterns.

To sum up, only the indefinite article *a*, *an* in the subject NP position was being realized at criterion or near categorical level. This is shown by pluses (+) whereas variable use was marked by crosses (x), and pre-systematic use by minuses (-). Errors in scaling patterns were circled. As regards the definite article *the*, overall performance remained either variable or pre-systematic. Finally, Table 5.25 shows that the performance on the *the*<sub>2</sub> was variable across levels in the subject NP environment; pre-systematic in the object NP environment; and again, variable in the PP environment, where performance was unpredictable.

## **5.4. Developmental Stages in the Use of Reference**

### **5.4.1. The Notion of Developmental Stages in SLA**

The notion of developmental stages evolved out of a major concern in IL studies in the late 70's, with how to characterise the linguistic and conversational context of IL performance (Long and Sato, 1984). This involves both the context IL speakers or users create for themselves, and the context created for them by their interlocutors. It also concerns the relationship between the development of particular sub-systems in the contexts of the wider IL grammar. Within the ZISA project study, Meisel, Clahsen and Pienemann (1981) investigated the word-order and certain movement rules in untutored migrant workers acquiring German as a Foreign Language. Meisel et al. (1981: 124) found that passage through the developmental stages of German word-order, temporarily, involved learners in deleting other elements such as V or object NP (over which one has to move) and subject or V (one of the categories to be inverted) which they had previously displayed in their ILs, apparently in an attempt to incorporate the new and more complex sets of movement rules required at each stage.

### **5.4.2. Determining Developmental Stages in the Use of Reference**

In Table 5.25, the variable performance on the article system was given, showing the probability for supplying the correct article in three environments across the five groups. Now we will look at particular individuals to examine, more accurately, the systematicity of the underlying trends. In other words, it is hypothesized that language development is not linear, but rather multi-dimensional. In addition, since language development is both dynamic and organic, changes occur not only along the temporal dimension but also by spreading towards more and more linguistic (and conversational) contexts. In

order to capture this kind of systematic change we selected a random sample of 71 out of 163 cases from all three levels of training, so that developmental stages could be determined according to the above assumptions.

In Table J (see Appendix), these individual subjects were subdivided into five broad developmental stages. These results show that the subjects performance was indeed highly variable and systematic at the same time. For instance, if we go to Stage I in Table 5.25 (i.e. the lowest stage), we can see that most subjects in that stage were performing poorly (i.e. almost never reached the 80% criterion). However, subjects no. 10, 115 and 11 (all at stage I) had already reached the  $\geq 80\%$  criterion for either the indefinite object NP, or the definite object NP; and yet <sup>the latter</sup> was the most difficult environment of all. If we then look up Stage V (the highest), we find that subject no. 165 who can be said to have mastered the whole article system at the  $\geq 80\%$  criterion, was still using the indefinite article in the PP position with only 40% accuracy (i.e. pre-systematic-like behaviour). One might be tempted to think that subject no. 165 belongs to the advanced group in terms of exposure or level of training; and that the other subjects mentioned earlier (no.s 10, 115 and 11) belong to less advanced levels. In fact, subject no. 165 is a school pupil. By comparison, subjects 10 and 11 belong to the teacher training college; and yet they still have not mastered even the indefinite article in the easiest environment, i.e. the subject NP position.

In order to be able to delineate these different stages more easily, we constructed a bimodal implicational table using pluses (+) and minuses (-), rather than a fully quantitative one. Table 5.26 gives the five developmental stages involving 71 subjects randomly selected from all three levels of training. The columns indicate the subjects' identification number and level (viz. school



pupils, university undergrads, and teacher trainees). The rows show the articles and linguistic environments. At the bottom of the table, the various cut-off points delineating the five stages are shown. Finally, pluses (+) are used to indicate categorical use (80-100 %); minuses (-) refer to variable use; and circles are used to indicate errors to the implicational patterns. An error in this particular instance was committed if one subject mastered the most difficult environment, viz object NP or PP, before the easiest, i.e. the subject NP position.

Table 5.26

**Developmental Stages Among 3 Articles Within 3  
Environments: Bimodal Implicational Patterns, N=71**

S.No.	A, AN			THE			THE2		
	AGT	OBJ	PREP	AGT	OBJ	PREP	AGT	OBJ	PREP
STAGE V (subject's level)									
165 (1)	+	+	-	+	+	+	+	+	+
74 (3)	+	+	+	+	+	+	-	-	⊕
160 (1)	+	+	+	+	+	+	-	-	⊕
STAGE IV									
157 (1)	+	+	+	+	+	-	-	⊕	-
6 (3)	+	+	+	+	+	+	+	+	-
41 (3)	+	+	+	+	+	+	+	-	-
135 (1)	+	+	+	+	+	+	-	-	⊕
7 (3)	+	+	-	+	+	+	+	-	+
147 (1)	+	+	-	+	+	+	+	-	+
91 (2)	+	+	-	+	+	+	+	-	-
96 (2)	+	+	-	+	+	+	-	-	⊕
141 (1)	+	+	-	+	+	+	-	-	⊕
2 (3)	+	+	+	-	-	⊕	-	-	⊕
80 (2)	+	+	+	-	-	⊕	-	-	⊕
66 (3)	+	+	-	+	-	+	-	-	⊕
95 (2)	+	+	-	-	-	⊕	+	-	+
155 (1)	+	+	-	-	-	⊕	+	-	+
144 (1)	+	-	+	+	+	+	+	-	+
162 (1)	+	-	-	+	+	+	+	+	+
5 (3)	+	-	-	-	⊕	⊕	+	+	+

## Key:

+ Categorical Use

- Variable Use

⊕ Error, i.e. if the Obj NP or PP environment is mastered before subject NP environment

TABLE 5.26 CONTINUED

S.No.	A, AN	THE	THE2
STAGE III			
95 (1)	+	+	+
77 (2)	+	+	+
118 (1)	+	+	+
43 (3)	+	+	+
142 (1)	+	+	+
134 (1)	+	+	+
93 (2)	+	+	+
45 (3)	+	+	+
18 (3)	+	+	+
37 (3)	+	+	+
159 (1)	+	+	+
87 (2)	+	+	+
68 (3)	+	+	+
166 (1)	+	+	+
79 (2)	+	+	+
139 (1)	+	+	+
145 (1)	+	+	+
114 (1)	+	+	+
133 (1)	+	+	+
126 (1)	+	+	+
72 (3)	⊕	⊕	+
112 (1)	-	-	+
103 (1)	⊕	-	-

## Key:

+ Categorical Use

- Variable Use

⊕ Error, i.e. if the obj NP or PP environment is mastered before subject NP environment

TABLE 5.26 CONTINUED

S.No.	A, AN			THE			THE2		
	AGT	OBJ	PREP	AGT	OBJ	PREP	AGT	OBJ	PREP
STAGE II									
8	(3)	+	+	-	-	-	-	-	-
22	(3)	+	+	-	-	-	-	-	-
9	(3)	+	+	-	-	-	-	-	-
32	(3)	+	+	-	-	-	-	-	-
27	(3)	+	+	-	-	-	-	-	-
16	(3)	+	-	+	-	-	-	-	-
97	(2)	+	-	+	-	-	-	-	-
163	(1)	+	-	-	-	-	⊕	-	-
125	(1)	+	-	-	-	-	⊕	-	-
123	(1)	+	-	-	-	-	-	-	-
140	(1)	+	-	-	-	-	-	-	-
33	(3)	+	-	-	-	-	-	-	-
60	(3)	+	-	-	-	-	-	-	-
21	(3)	+	-	-	-	-	-	-	-
13	(3)	+	-	-	-	-	-	-	-
153	(1)	+	-	-	-	-	-	-	-
113	(1)	+	-	-	-	-	-	-	-
129	(1)	+	-	-	-	-	-	-	-
150	(1)	+	-	-	-	-	-	-	-
158	(1)	-	⊕	⊕	-	-	-	-	-
167	(1)	-	⊕	⊕	-	-	-	-	-
STAGE I									
11	(3)	-	-	-	-	-	-	⊕	-
90	(2)	-	-	-	-	-	⊕	-	-
115	(1)	-	-	-	-	⊕	-	-	-
10	(3)	-	⊕	-	-	-	-	-	-
92	(2)	-	-	-	-	-	-	-	-
138	(1)	-	-	-	-	-	-	-	-
146	(1)	-	-	-	-	-	-	-	-

## Key:

+ Categorical Use

- Variable Use

⊕ Error, i.e. if the obj NP or PP environment is mastered before subject NP environment

First of all, Stage I includes subjects who have not yet mastered even the easiest article (i.e. *a, an*) in the most favoured environment (i.e. the subject NP). Thus, there are minuses (-) almost everywhere in the table. To this extent, these subjects' performance is variable and pre-systemic; although some give the impression of having mastered (i.e. free variability) some of the more difficult forms and environments. As shown in sections 5.1 and 5.2 (by Scheffe tests and Scalogram analysis), the three articles and three environments in which they occur are intertwined and implicationally related; in such a way that mastery of *the*<sub>2</sub> would also imply that of *the* and *a, an*. In addition, mastery of the object NP environment would entail that of the subject NP, but not necessarily that of the PP environment (i.e. Subj. NP  $\geq$  PP  $\geq$  Obj. NP). All subjects are, therefore expected to conform to this pattern. To this effect, no subject could be said to have mastered systematically the definite article *the*<sub>2</sub> before the indefinite article *a, an*, or the object NP environment before the subject NP position. Whenever this occurs, we regarded this kind of free variability as random i.e. a violation of the systematic implicational pattern in the data (i.e. see circled + marks in the table).

Secondly, at Stage II, almost everyone is using the indefinite article at criterion level in the subject NP environment. Another characteristic of the learners' IL behaviour, at this stage, is to use the indefinite article categorically in the subject NP position as well as in either the object NP or the PP environment. However, subjects who seem to have mastered the set of object NP and PP environments before the Subject NP position are regarded as violating the pattern (e.g. nos. 158 and 167). At this stage, both the definite article *the* and *the*<sub>2</sub> are not being used systematically yet. Still, subjects 163 and 125 were violating this pattern since they are using *the* in the PP position at  $\geq$  80 criterion.

Thirdly, at Stage III, the most noticeable fact is that while the pattern for the use of the indefinite article remains largely the same as at Stage II, the definite article *the* is rapidly catching up with the indefinite article *a, an*. Thus we have subjects 159 and 87 who have mastered fully the definite article *the* in all three environments; but have only mastered the subject NP position in relation to the indefinite article. In addition, regarding the other definite article *the<sub>2</sub>*, there are the first signs of consistent use of this structure, although performance is highly variable. Towards the end of Stage III, we find for the first time, subjects who have by then mastered the indefinite article completely, i.e. in all three environments. These are subjects 95 and 77; and yet, subject 77 for instance is still in the pre-systematic Stage I as regards the definite article *the* and *the<sub>2</sub>*.

Fourth, at Stage IV, more and more subjects are now using correctly (or indeed concentrating more on) the definite article *the<sub>2</sub>*. Two subjects (viz. nos. 162 and 5) even appear to have reached categorical level in all three environments for the *the<sub>2</sub>*, which turns out to be exceptional as far as the whole sample was concerned. As expected, more systematic use is recorded in relation to the indefinite *a* and definite *the* articles. At Stage IV, 6 subjects (out of 17, i.e. 35%) are now using the indefinite article systematically in all three environments, compared with 7 out of 17 (i.e. 41%) who are using the definite *the* at the  $\geq 80\%$  criterion. These subjects are nos. 157, 6, 41, 135, 2 and 80 for the indefinite article *a, an*; and nos. 7, 147, 91, 96, 141, 144 and 162 for the definite article *the*. In addition there is an interesting feature about the type of change which is taking place. For the first time we find subjects (e.g. no. 162) who are now able to distinguish between the two subcategories of the definite article, viz. *the* and *the<sub>2</sub>* in all three environments. His performance can be referred to as systematic, but still variable, since he has to learn how to use

the indefinite article correctly in the object NP and PP environments.

Fifth, Stage V, comprised those few subjects who can use systematically both the indefinite article *a, an* and definite article *the*. Only three subjects are found to be in this stage, viz. nos. 165, 74, 160. Leaving *the<sub>2</sub>* article aside, it can be seen that these subjects have mastered fully the two main subcategories of the article system, although no. 165 was found to be still grappling with the indefinite article in the PP environment where his performance appears to be pre-systematic. Most analyses of the article system usually limit themselves to these two forms (i.e. *a, an* and *the*). However, we believe that complete mastery of the article system would entail that the learners become able to make the further fine-grained distinction between the definite article whose referent is linguistically recoverable (i.e. *the*) and that whose referent is not linguistically identified (i.e. *the<sub>2</sub>*). If we look at Stage V subjects' performance, we notice that indeed this particular distinction between *the* and *the<sub>2</sub>*, based on a form to function analysis of the definite article, remains a problem for these students.

To sum up, the delineation of the developmental stages is, of necessity, somewhat arbitrary; but it can help shed light on the systematic variability in IL behaviour, to enable us to understand the learning process better. The number of characteristics of the stages might vary as the learners proceed on to acquire more facts about the target language. Thus, the stages may depend on the type of data from particular learners, although one might expect generalizable commonalities among adult learners in similar learning conditions.

## 5.5. Discussion

Our investigation of the definite and indefinite reference among these Zairean learners of English shows that the variability in the use of the forms *the* and *a* before their functional mastery is not random. Our results suggest that

this variability is systematic and can be described in terms of the changing rules governing semantic features of the noun phrase. Our approach has been a narrow and intensive one, concentrating on a small but rich aspect of semantics rather than a large number of related aspects of reference. We hope that the narrowness and intensive nature of our investigation in this area has enabled us to provide a sharper picture of these learners' abilities to use definite and indefinite reference at particular stages of their developing grammars. On the basis of our quantitative analysis (by ANOVA and Scheffe tests) we have found no significant differences between these students' performance on the use of reference although they happen to belong to three typically different levels of proficiency (viz. high school, university and teachers' college). This apparent lack of development may be due to the fact that definite and indefinite reference represent a highly complex subsystem, and thus, may not be mastered fully until very late in the learning process.

When we examined our qualitative results (mainly based on individual IL behaviour), we were then able to delineate a number of developmental stages. These five stages have helped us demonstrate ways in which IL users gradually establish appropriate relationships between the articles (i.e. forms) in the contexts of wider IL subsystem of definite and indefinite reference. For instance, we have been able to establish that the indefinite article *a* is being learned at a faster rate than *the*. However, in certain stages (e.g. Stage III) the learners had to neutralize (or slow down) the development of *a* in certain linguistic environments (e.g. object NP) so that progress can be achieved as regards *the* in either the subject NP or object NP environments. In addition, the subject NP environment appears to be easier than either the object or prepositional phrase. This finding lends support to both Givon (1984) and Huebner's (1983) proposal that the presupposed - asserted word-order and the



agent/experiencer verb-object word-order are compatible.

Another interesting finding involving definite reference is the fine-grained distinction that we have established between *the* (i.e. with referent linguistically present in the text) and *the*<sub>2</sub> (i.e. with referent non-linguistically identifiable). This distinction seemed to represent the complex notion of specificity (e.g. 'entailment', Karttunen 1968) discussed in the theoretical section of this investigation. Most analyses to date have usually lumped together these two functions of definite reference. And yet, in our developmental Stage V, we find that even learners who could be credited with the near native mastery of the semantics of the articles (i.e. *a* and *the*), are still having great difficulty with the finer distinction involving *the*<sub>2</sub>.

Moreover, one of the issues that have been suggested is whether the so-called zero article *0* could be appropriately investigated within our framework (i.e. using production data). We believe that it is misleading to regard instances of omission of an article as evidence that the function of properly referring within discourse has been achieved. Since omission of an article is ambiguous (involving knowledge or no-knowledge of the appropriate form/function) we believe that some other elicitation techniques (e.g. interviewing the subjects as to what they intended to say) would probably be more appropriate. Finally, although no theoretical claims have been entertained in relation to such issues as whether the semantic restructuring terminates earlier or later than its syntactic counterpart, our results indicate that at the beginning stages, syntactic forms usually appear or are often used independently of their meaning. It is only gradually that the relationships between form and function seem to be mapped. On this account, it would seem that in SLA, syntax proceeds faster than semantics; i.e. in a similar

manner that learning putatively takes place at a faster rate than acquisition (in Krashen's 1981 terms).

## CHAPTER 6

## RESULTS OF THE INVESTIGATION OF THE COMPLEMENT

## SENTENCES AND UNIVERSAL TOPICALITY HIERARCHIES

**6.1. Group Trends**

The investigation of the infinitival complement focuses on the way in which the subjects would interpret complex sentences including one main clause with the *promise, ask, tell* verb types, and one embedded subordinate clause in which the co-referential subject has been deleted. These types of constructions are called infinitival complement clauses, in syntactic terms. As with the use of definite and indefinite reference, variability in the subjects' performance was analysed: first, by using a modified version of Andersen's (1978) Group Range Method. Secondly, an analysis of variance (ANOVA) with Scheffe tests was conducted. Thirdly, implicational scales were built both manually and by the Guttman Scale Computer Program.

**6.1.1. Group Range Variability**

Six Group ranges were established in order to show the kind of variability within each group or levels of training. As can be seen from Table 6.1 to 6.3, the six ranges were defined in terms of percentage of correct responses over the total number of items (there were 36 sentences to interpret, thus, 36 items in this task).

Table 6.1

## Group Range Variability of Promise Verbs

Range of % correct	Level1 High school	Level2 Undergrad.	Level3 Teacher Tr.	Total
0 - 20 %	12 (18.2%)	1 ( 4.8%)	19 (25.0%)	32 (19.6%)
21 - 40 %	12 (18.2%)	1 ( 4.8%)	13 (17.1%)	26 (16.0%)
41 - 60 %	15 (22.7%)	5 (23.8%)	13 (17.1%)	33 (20.2%)
61 - 80 %	9 (13.6%)	7 (33.3%)	8 (10.5%)	35 (21.5%)
81 - 99 %	8 (12.1%)	2 ( 9.5%)	5 ( 6.6%)	15 ( 9.2%)
100	10 (15.2%)	5 (23.8%)	7 ( 9.9%)	22 (13.5%)
TOTAL	66	21	76	163
Number > 80%	18 (27.3%)	7 (33.3%)	12 (15.8%)	37 (22.7%)

Table 6.2

## Group Variability of Ask Verbs

Range of % correct	Level1 High school	Level2 Undergrad.	Level3 Teacher Tr.	Total
0 - 20 %	1 ( 1.5%)	0 ( 0%)	0 ( 0%)	1 ( 0.6%)
21 - 40 %	5 ( 7.6%)	0 ( 0%)	5 ( 6.6%)	10 ( 6.1%)
41 - 60 %	14 (21.2%)	4 (19.0%)	17 (22.4%)	35 (21.5%)
61 - 80 %	12 (18.2%)	9 (42.9%)	18 (23.7%)	39 (23.9%)
81 - 99 %	12 (18.2%)	7 (33.3%)	11 (14.5%)	30 (18.4%)
100%	22 (33.3%)	1 ( 4.8%)	25 (32.9%)	48 (29.4%)
TOTAL	66	21	76	163
Number > 80%	34 (51.5%)	8 (38.1%)	36 (47.4%)	78 (47.9%)

Table 6.3

## Group Variability of Tell Verbs

Range of % correct	Level1 High school	Level2 Undergrad.	Level3 Teacher Tr.	Total
0 - 20 %	1 ( 1.5%)	0 ( 0%)	0 ( 0%)	1 ( 0.6%)
21 - 40 %	7 (10.6%)	1 ( 4.8%)	2 ( 2.6%)	10 ( 6.1%)
41 - 60 %	14 (21.2%)	4 (19.0%)	19 (25.0%)	37 (22.7%)
61 - 80 %	15 (22.7%)	6 (28.6%)	20 (26.3%)	41 (25.2%)
81 - 99 %	11 (16.7%)	5 (23.8%)	13 (17.1%)	29 (17.8%)
100%	18 (27.3%)	5 (23.8%)	22 (28.9%)	45 (27.6%)
TOTAL	66	21	76	163
Number > 80%	29 (43.9%)	10 (47.6%)	35 (46.1%)	74 (45.4%)

If we consider the *promise* verbs first, it can be seen that nearly 20% of all subjects are still in the 0-20%, the lowest range. However, almost the same proportion of subjects are found in three other ranges as well, viz. 20.2% in the 41-60% range; 21.5% in the 61-80% range; and 22.7% in the top 81-100% range. These results suggest that the subjects' performance on this particular verb type is far from being homogeneous. Let us then look at each level in particular.

At level 1 (i.e. the school pupils) the greatest proportion of the subjects are

in the top range (27.3%). This pattern is found at level 2 (i.e. the university students), where as many as 33.3 % of the respondents are in the 81-100 % range. However, for the advanced level, the largest proportion of respondents are found within the 0-20 % range (viz. 25 %).

Secondly, let us examine the performance on the *ask* verbs. In Table 6.2, it can be seen that almost half of the total number of subjects are found within the highest range; there are as many as 47.9 %. Meanwhile, 23.9% of the total number of the subjects are found in the 61-80 % range and 21.5 % in the 41-60 % range. Clearly the performance on this verb type was much better, since only 6.7 % of the subjects were found in the (lowest) 0-20 % range compared with 19.6 % in this range for the *promise* verbs. If we consider the different levels separately, we find that (in Table 6.2) more than one half of the school pupils (51.5%) are in the top range. The pupils are doing relatively better than any other level, as regards performance in the 81-100 % range (at level 2, there are 38.1 % in this top range, whereas at level 3 we find 47.4 %). But, still at level 1, notable proportions of subjects are also found in the intermediate 61-80 % and 41-60 % ranges (viz. 18.2 % and 21.2 % respectively).

As far as the university students (level 2) are concerned, 38.1 % of the respondents are found in the 81-100 % range. However, the greatest proportion of the respondents at this level are still within the intermediate 61-80 % range. Only 19 % are found in the 41-60 % range; but none at all in the lower ranges (the 0-20 % and 21-40 % group ranges). As for the teacher-trainees, as at the other two levels, the largest proportion of respondents were found in the top 81-100 % range (i.e. 47 %). But there were also notable proportions in the intermediate ranges too (i.e. 23.7 % and 22.4 % in ranges 61-80 and 41-60, respectively). Only 6.6 % of the respondents are found in the two lowest group

ranges.

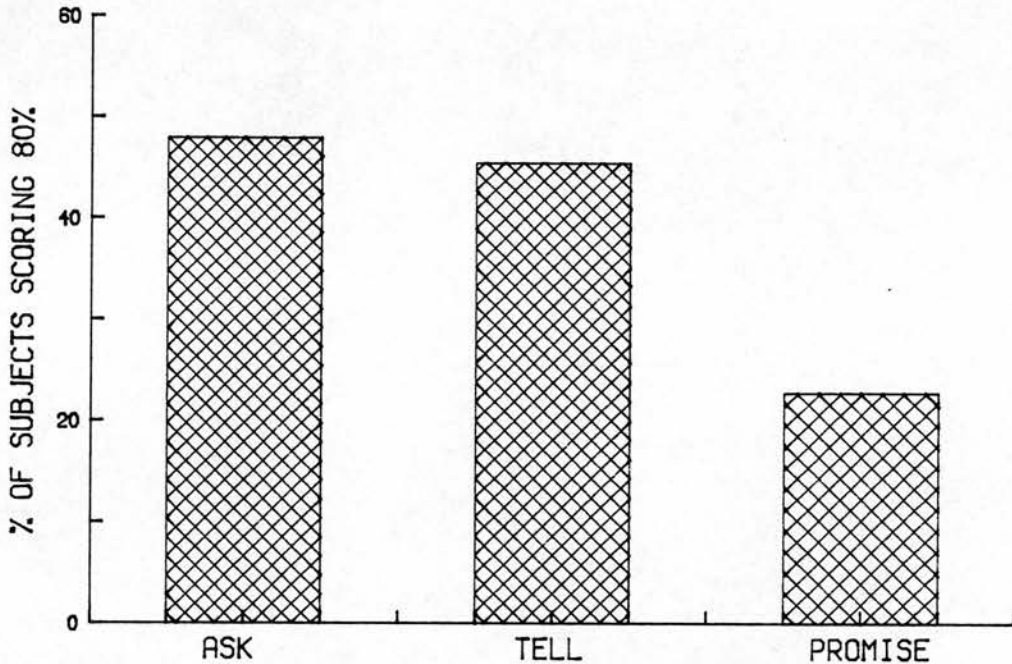
In Table 6.3, the group range results for the *tell* verbs are reported. As with the *ask* verbs, we find that a greater proportion of the respondents are in the 81–100 % group range than in any of the other four ranges. By looking at the overall results for all levels combined, it appears that as many as 45.4 % are in the top range (80–100 %); 25.2 % were in the 60–80 % and 22.7 % in the 41–60 % range. Only 6.7 % of the respondents were in ranges 21–40 and 0–20 %.

When we look at the different levels separately, we realize that basically the pattern remained the same within each level as with all combined together. This pattern indicates that there were more respondents in the 81–100 % range than in any of the lower group ranges. These results show that the respondents were highly accurate in interpreting the *tell* verb construction. The percentages of subjects giving correct responses within the 81–100 % range are as follows: for school pupils, 43.9 %; for university undergraduates 47.6 %, and for teacher–trainees 46.1%. Again, as in the case of the *ask* verbs, substantial proportions of subjects were also found in the intermediate ranges 41–60 % and 60–80 %; whereas only a comparably smaller proportion of the subjects were in ranges 0–20 and 21–40. These findings are summarized in Figure 6.1. This histogram shows the percentage of respondents scoring 80 % or more for each verb type. It is clear that the performance of all groups of subjects combined was better as regards *ask* than *tell* verbs. The *promise* sentences proved more difficult to interpret. The difficulty may be due to the fact often pointed out in English that the *promise* verb type does not follow the general pattern.



Figure 6.1

Histogram for three verbs



The majority of the subjects (77.3 %) believed that the object of the main clause was always the subject of the embedded clause. However, if we examine the respondents' performance on both *ask* and *tell*, we find that it was not so obvious (on basis of this histogram) which of the two verbs was causing greater difficulty. In view of this fact, we have to conduct analysis of variance to be able to establish how big the differences might be between: 1) verb types; 2) levels of training in English; and 3) features along the dimension of topicality hierarchy.

#### 6.1.2. Analysis of Variance: ANOVA and Scheffe Tests

In the following paragraphs the results of the ANOVA and Scheffe procedures are given. The ANOVA procedure was conducted by BMDP 2PV computer program for ANOVA with Repeated Measures and Covariance. The

Scheffe tests were carried out following Winer (1971: 198) The experimental design involved the following variables:

1. Three verb types (i.e. *promise*, *tell*, *ask*);
2. Three subject types based on the semantic features 'human', 'inanimate' or 'animate';
3. Three object types marked for 'human', 'inanimate' or 'animate';
4. Three levels of training (i.e. pupils, undergrads, teacher trainees)

It was hypothesized that comprehension of the *promise*, *tell*, *ask* constructions would depend upon both the syntax of the target language and semantic features or role relationships of the first and second noun phrase in the sentence. It was further assumed that of all three types of constructions, those with *tell* verbs would be easier to grasp since they follow the general pattern in English, which requires that the object of the main clause becomes the subject of the subordinate. In addition, since *promise* verbs belong to an exceptional class, it was hypothesized that these constructions would lead to some confusion, and thus would be more difficult to interpret. As regards constructions with *ask* verbs, it was assumed that these would cause some difficulty since they can be ambiguous to interpret. For instance, in the following sentences it might be difficult to decide who should have the shower:

- 1.) The mother asked the child to have a shower.
- 2.) The child asked the mother to have a shower.

Syntactically speaking, in (1) the child should have the shower, whereas in (2) the mother should. However, *ask* is ambiguous semantically; that is, it can imply 'X requests Y to do something' or 'X asks for Y's permission'. Taking into account the pragmatic role relationships between the mother and the child, one

may arrive at at least two interpretations for each of the sentences (1) and (2) above, depending on whether *ask* is a request or a search for approval. Because of this possibility of an ambiguous interpretation, the *ask* constructions were assumed to cause problems, just in case the second interpretation (Asking for permission) was preferred. Table 6.4 gives the results of the analysis of variance.

Table 6.4

ANOVA Results for 3 Verbs, 3 Subjects and 3 Object Types  
by Levels

SOURCE	SS	d.f.	MS	F	P
Main effects:					
Levels	2.101	2	1.050	1.12	0.328
Verbs	25.585	2	12.292	25.50	0.000**
Subjects	0.082	2	0.041	0.32	0.724
Objects	0.276	2	0.138	1.12	0.329
2-Way Interaction:					
V x Level	5.736	4	1.434	2.98	0.02*
S x Level	1.475	4	0.369	2.89	0.02*
V x Subject	11.004	4	2.751	16.34	0.000**
V x Object	11.881	4	2.970	21.48	0.000**
S x Object	3.395	4	0.849	6.05	0.000**
3-Way Interactions:					
V. O. L.	2.914	8	0.364	2.63	0.007**
V. S. L.	7.499	8	0.937	6.66	0.000**
4-Way Interactions:					
V. S. O. L.	4.738	16	0.296	2.10	0.007**
** p < .01   * p < .05					

The results of ANOVA show that the main effect for verb type was highly significant ( $F = 25.5$ ;  $df = 2$ ;  $p < .0001$ ). This indicates that the respondents' performance varied very significantly depending on the verb type. No statistical significance was found concerning the other main effects (i.e. levels of training, subject type or object type). Appropriate Scheffe tests were conducted to determine which verb type was causing more difficulty. These results are shown in Table 6.5. They reveal that *promise* verbs were the most difficult, whereas there was no difference at all between the performance on *tell* and *ask*

(p &lt; .01).

Table 6.5

## Scheffe Tests for Comparison between Three Verbs

	Promise	Ask	Tell
mean=	762.84	1085.58	1085.58
Promise	-	840.42**	840.42**
** p < .01 t'crit = 80.83			

The next step in our investigation is to examine interaction effect between verb type and level of training. In Table 6.4, we can see that there was a significant interaction effect between verbs and levels of training ( $F = 2.98$ ;  $df 4$ ;  $p < .02$ ). This suggests that some verbs were more difficult for the respondents at a certain level of training than others. In view of this finding we have to examine which verb might be causing difficulty for which level of respondents. These Scheffe tests are given in Table 6.6. These results reveal that as expected, the teacher trainees (level 3) did significantly better than the other two levels as regard *tell* and *promise* verbs, followed by the school pupils; and conversely, the university students performed worse than the other two levels on both *tell* and *promise* verbs. It is further confirmed that for all levels, the *promise* constructions are significantly more difficult than either *tell* or *ask*. In relation to the *ask* constructions, the trend is the same; that is, the teacher-trainees are in the first position; then, the pupils come second; and the undergrads in third position. All these results are statistically significant ( $p < .01$  or  $p < .05$ ). Another finding revealed by these results (in Table 6.6) is that, in some combinations of verb level, there was no significant difference between the most advanced students (i.e. teacher-trainees) and the least advanced (i.e. the

Table 6.6

**Scheffe Tests for Comparisons between Verbs and Three  
Levels of Training in English**

	1 Promise Ugrads	2 Ask Ugrads	3 Tell Ugrads	4 Promise School	5 Promise Pgrads	6 Tell School	7 Ask School	8 Ask Pgrads	9 Tell Pgrads
mean=	130.41	139.86	143.64	314.82	321.48	427.68	445.50	513.00	526.68
1. Promise x UG	--			** 184.41	** 191.07	** 297.27	** 315.09	** 382.59	** 396.27
2. Ask x UG		--		** 174.96	** 181.62	** 287.82	** 305.64	** 373.14	** 386.82
3. Tell x UG			--	** 171.18	** 177.84	** 284.04	** 301.86	** 369.36	** 383.04
4. Promise x Sch				--		** 112.86	** 130.68	** 198.18	** 211.86
5. Promise x PG					--	** 106.20	** 124.02	** 191.52	** 205.20
6. Tell x Sch						--	** 85.32	** 99.00	** 99.00
7. Ask x Sch							--	*	81.18
** P < .01									
* P < .05									

pupils). This pattern was found with both verbs *promise* and *ask*, as can be seen by looking at the cross-section of row no 4 (*promise* x school) with column 5 (*promise* x postgrads). In connection with *ask*, a similar result can be observed by inspecting the crossing point of row 7 and column 8 (the blank shows that the means involved in the comparison are not different enough to reach any significance). Recall that one of the hypotheses under investigation is that as the students' level of proficiency increases, their ability to comprehend these complex constructions would also increase. Although this assumption is borne out in many instances, these results show that in other instances the pupils' and teacher-trainees' performance does not differ. There was also a significant 2-way interaction effect for verbs and subject noun phrase ( $F = 16.34$ ;  $df = 4$ ;  $p < .0001$ ) as well as for verbs and object noun phrase ( $F = 21.48$ ;  $df = 4$ ;  $p < .0001$ ). This suggests that the difficulty in interpreting these sentences was dependent on whether either the subject or object of the main clause had the semantic features of + Human, + Animate, or + Inanimate. Various a posteriori tests were conducted to identify the combinations of verbs and subject or verbs and object which reached statistical significance. These Scheffe tests are summarised in Tables K, L, and M (see appendix I). They indicate that in most cases a particular sentence type became easier to interpret if its subject, together with its object NPs, involved certain semantic features rather than others. However, as regards the *promise* constructions, this variation of semantic features made little difference (i.e. the *promise* sentences were causing difficulty no matter what features the subject or object of the main clause had). Next, we examine the 3-way interaction effect between verb, subject and object combined together. Indeed we found (Table 6.4) that such an interaction effect was significant ( $F = 6.66$ ;  $df = 8$ ;  $p < .0001$ ). These results suggest that the combination of a particular verb with certain NPs as subject

and certain NPs as object did affect significantly the respondents' interpretation of certain sentences.

Appropriate Scheffe tests were again carried out to identify significant differences between pairs of combinations of verb-subject-object. These results are given in Table 6.7. They suggest that, as Gass (1984) has pointed out, the acquisition (or mastery) of the complex sentences involves a combination of both syntactic forms and semantic features. Gass (op. cit.) argues that universals, seen in the role of the Topicality Hierarchy, interact with the language specific facts of English. In other words, the syntactic mastery of *promise, tell, ask* verbs was differentially affected by different hierarchical role relationships, involving human, inanimate, and animate. Finally, although Table 6.4 showed a significant 4-way interaction effect, (i.e. verb, subject, object and level), we decided not to proceed with the Scheffe procedure, which would involve testing for differences between no less than 81 possible comparisons. Although the Scheffe procedure would yield very detailed results, their contrastive strength or clarity would be greatly reduced. We shall then deal with further scrutiny of data using qualitative analysis, referred to as 'Learner Language Behaviour' (section 6.3). First, however, we will examine the main implicational patterns of grammatical accuracy in the learners' performance on the complement sentences.

### 6.1.3. Scalogram Analysis

In the last two subsections we looked at group trends variability within the whole range of data, using the group range method and ANOVA-Scheffe tests. It appeared that in general, *promise* constructions were causing more difficulty, whereas both *tell* and *ask* were easier to interpret across the three levels of training under investigation. The next thing we do is to see if the same kind of



Table 6.7

Scheffe Tests for a 3-way interaction between Verb, Subject and Object

	PROMISE Human S x Anim. O	PROMISE Inan. S x Hum. O	PROMISE Inan. S x Anim. O	PROMISE Anim. S x Anim. O	PROMISE Anim. S x Hum. O	ASK Hum. S x Inan. O	PROMISE Anim. S x Inan. O	TELL Anim. S x Inan. O	PROMISE Inan. S x Inan. O	PROMISE Hum. S x Inan. O	TELL Inan. S x Inan. O	ASK Anim. S x Inan. O	ASK Hum. S x Hum. O	TELL Inan. S x Hum. O	TELL Hum. S x Anim. O	ASK Hum. S x Anim. O	TELL Inan. S x Anim. O	ASK Anim. S x Hum. O	ASK Inan. S x Inan. O	TELL Anim. S x Hum. O	ASK Anim. S x Anim. O	TELL Anim. S x Anim. O	TELL Hum. S x Hum. O	ASK Inan. S x Anim. O	TELL Hum. S x Inan. O	ASK Inan. S x Hum. O				
T	T.1	T.2	T.3	T.4	T.5	T.6	T.7	T.8	T.9	T.10	T.11	T.12	T.13	T.14	T.15	T.16	T.17	T.18	T.19	T.20	T.21	T.22	T.23	T.24	T.25	T.26	T.27			
N	69	73	74	78	79	95	97	98	99	100	105	108	110	111	114	114	116	116	119	122	123	124	140	142	149	150	155			
T.1	--											41	42	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**		
T.2		--												41	41	42	43	46	49	50	51	67	69	76	77	82	**	**	**	
T.3			--													42	42	45	48	49	50	56	58	75	76	81	**	**	**	
T.4				--													41	44	45	46	62	64	71	72	77	**	**	**	**	
T.5					--													43	44	45	61	63	70	71	76	**	**	**	**	
T.6						--															45	47	54	55	60	**	**	**	**	
T.7							--														42	45	52	53	58	**	**	**	**	
T.8								--														42.2	44.2	51.2	52.2	57.2	**	**	**	**
T.9									--													41	43	50	51	56	**	**	**	**
T.10										--													42	49	50	55	**	**	**	**
T.11											--													44	45	50	**	**	**	**
T.12												--												41	42	47	**	**	**	**
T.13													--																	45
T.14														--																44
T.15															--															41
T.16																	--													41

\*\* P < .01  
\* P < .05  
+ P < .10

systematic variability could be established when considering only those subjects who could be credited with the near categorical performance, i.e. those who had reached the > 80% criterion. In order to do this, we shall conduct scalogram analysis using the Guttman Scale SPSS subprogram. The cut-off point is to be set at the > 80% criterion. More precisely, this criterion is considered met if the respondent scores seven items correctly out of a maximum of nine items per verb type (the exact proportion is thus 0.78, which is the nearest one could get to 0.80; also this figure could normally be rounded to one decimal place, which equals 0.8). The Guttman scale results are given in Table 6.8. It is thus established that in fact the order of difficulty between these verb types was *tell* > *ask* > *promise* (coef. of reproducibility = 0.88; coef of scalability = 0.70).

Table 6.8

## Guttman Scale for 3 Verb Types

ITEM	PROMISE		ASK		TELL		TOTAL
RESPONSE	0	1	0	1	0	1	
SCALE	-----ERR-----		-----ERR-----		-----ERR-----		
3	0	34	0	34	0	34	34
	---ERR						
2	44	10	5	49	5	49	54
			-----ERR				
1	22	11	25	8	19	14	33
					---ERR		
0	42	0	42	0	42	0	42
SUMS	108	55	72	91	66	97	163
PCTS	66	34	44	56	40	60	
ERRORS	0	21	5	8	24	0	58

163 cases were processed  
 0 (or 0.0 PCT were missing)

## STATISTICS

1. Coefficient of Reproducibility = 0.88
2. Minimum Marginal Reproducibility = 0.61
3. Percent Improvement = 0.28
4. Coefficient of Scalability = 0.70

## YULE'S Q

	PROMISE	TELL	ASK
Promise	1.00	0.36	0.45
Tell	0.36	1.00	0.89
Ask	0.45	0.89	1.00

These results indicate that there is a valid scale for the above ordering to represent the most favoured implicational pattern among the three levels of training combined together. We feel totally confident in the significance of this finding since it has already been established by ANOVA- Scheffe tests when

we were dealing with the whole range of variability in the data earlier on.

In addition, the Guttman scale (Table 6.8) shows high correlation coefficients (Yule's Q) between *tell* and *ask* (viz. 0.89, which accounts for 79 % of common variance between the performance in the two verbs). By comparison, there was little in common between either *promise* and *tell* (viz. 0.36, or 13 % only of variance overlap) or *promise* and *ask* (viz. 0.45, or 20 % of the shared variance). After establishing this systematic pattern between the three verb types (as reflected by the respondents' performance in general), we shall conduct another implicational scale to assess what effect the subject and the object NPs might have on the correct interpretation of the embedded clause. This was a combination of features such as 'human' / 'human'; 'animate' / 'animate'; 'inanimate' / 'inanimate', as well as all possible combinations between or across these features.

The results of this implicational scale are given in Table N (see appendix I). These findings reveal notable values for the coefficient of reproducibility; but these results are not high enough to yield a valid scale (coef rep. = 0.81, coef. scal. = 0.33). So it is not possible to predict precisely on the basis of these results which combination of noun phrase features are creating more difficulty for sentence interpretation. On the other hand, when we compare these results in Table N with those obtained by the Scheffe procedure (in Table M), we find that it is possible to detect a certain tendency for particular combinations of features to appear at the extreme ends of the continuum (in both tables). For both Scheffe tests and Guttman scales, there appears to be some similarity in relation to the five right-most combinations or the five left-most (yielding a 10 point scale of variable difficulty). Those combinations of subject/ object NPs that are more likely to cause difficulty are as follows (only three out of five

from each extreme of the scale are given):

**Table 6.9**

**Three left-most Combinations of Features for NP<sub>1</sub> and NP<sub>2</sub>**

TEST TYPE	More difficult	Ave. Difficulty	Less Difficult
Guttman	Human x Anim.	Anim. x Human	Inan. x Anim.
Scheffe	Human x Anim.	Anim. x Inan. Anim. x Human	Inan. x Inan. Inan. x Anim.

The orderings given in Table 6.9 indicate that the most difficult combination of subject NP and object NP was when these two had the features of animacy (i.e. human/animate, and animate/human). At first, this seems surprising since the combination human/animate, for instance converges with the principle of minimal distance which accounts for the main pattern in English. However, by looking at Table 6.10 we notice that this combination is particularly difficult with the verb *promise* only. In Table 6.10 we have the three right-most combinations of NPs features which were less difficult.

**Table 6.10**

**Three Right-most Combinations of Features for NP<sub>1</sub> and NP<sub>2</sub>**

TEST TYPE	Less easy	Average	Easiest
Guttman Scale	Human x Inan.	Anim. x Anim.	Human x Human
Scheffe Test	Anim. x Anim.	Human x Human	Human x Inan.

The orderings in Table 6.10 suggest that the easiest combination of features is when both NPs (i.e. subject and object) are 'human'; even with the *promise* verbs (in Table 6.7), this trend seems to be borne out. This again seems an

indirect indication that differential or conflicting role relationship between NP<sub>1</sub> and NP<sub>2</sub> does affect the interpretation of these complex sentences. Thus when both NPs are 'human' we have a neutral situation where there is no semantic conflict since both NPs can be agents of an action verbs. In this case, the general pattern of English may apply. Briefly, the most likely order (for group trend) of difficulty, from easiest to most difficult appeared to be as follows:

Table 6.11

**Combinations of Features in Increasing Order of Difficulty**

TEST TYPE	EASIEST		MOST DIFFICULT	
GUTTMAN SCALE	Human S	An. S	An. S	Human S
	with	with	with	with
	Human O	An. O	Human O	An. O
SCHEFFE TESTS	Human S	Human S	An. S	Human S
	with	with	with	with
	Inan. O	Human O	Inan. O	An. O
			or Human O	

In the next subsection we shall examine how individual subjects were dealing with these complex constructions.

## 6.2. Individual Trends in the Performance on Complement Structures

### 6.2.1. Implicational Scales for Verb Types

In the previous section it was found that the three verb types were likely to be learned in the following order: *tell* > *ask* > *promise* when all groups were combined together. Now we are going to examine to what extent this implicational pattern may apply for particular individuals at different levels of training. In Tables 6.12, 6.13 and 6.14 the implicational scales for these three

levels are given.

Table 6.12

**Implicational scale for Individuals' Realization of  
Three Verbs [66 Pupils]**

S.No.	Ask	Tell	Promise
133	+	+	+
167	+	+	+
118	+	+	+
141	+	+	+
142	+	+	+
150	+	+	+
145	+	+	+
128	+	+	+
127	+	+	+
160	+	+	+
143	+	+	+
108	+	+	+
130	+	+	-
166	+	+	-
119	+	+	-
138	+	+	-
139	+	+	-
106	+	+	-
144	+	+	-
126	+	+	-
165	+	+	-
158	+	+	-
153	+	+	-
159	+	+	-
123	+	+	-
125	+	+	-
131	-	+	⊕
163	+	-	⊕
134	+	-	⊕
151	+	-	⊕
112	-	+	-
103	-	+	-
116	+	-	-
111	+	-	-
115	+	-	-
156	+	-	-
109	+	-	-
155	-	-	⊕
Correct	34	29	18 = 81
Errors	5	3	6 = 14

Coef. Rep. = 0.93

Coef. Scal. = 0.88



For level 1 (i.e pupils) the best or preferred order appears to be as follows: *ask*

> *tell* > *promise* (coef. rep. = 0.93; coef scal = 0.88).

Table 6.13

**Implicational scale for Individuals' Realization of  
Three Verbs [21 Undergrads]**

S.No.	Ask	Tell	Promise
85	+	+	+
87	+	+	+
77	+	+	-
78	+	+	-
81	+	+	-
89	-	+	⊕
94	-	+	-
91	-	+	-
90	+	-	-
96	+	-	-
83	+	-	-
86	+	-	-
88	+	-	-
95	-	-	⊕
80	-	-	⊕
97	-	-	⊕
79	-	-	-
82	-	-	-
84	-	-	-
2	-	-	-
93	-	-	-
Correct	10	8	6 = 24
Errors	6	0	4 = 10

Coef. Rep = 0.84

Coef. Scal. = 0.74

As for level 2 (i.e undergrads), the best order was somewhat different, viz. *tell* > *ask* > *promise* (coef. rep. 0.84; coef. scal. =0.74). In addition for level 3 (i.e. teacher trainees) the most favoured order was identical to that of level 1 viz. *ask* > *tell* > *promise* (coef. rep. = 0.92; coef. scal. = 0.88).

Table 6.14

Implicational Scale for Individuals' Realization of  
Three Verbs [76 Postgrads]

S.No.	Ask	Tell	Promise
66	+	+	+
68	+	+	+
49	+	+	+
56	+	+	+
70	+	+	+
10	+	+	+
2	+	+	-
5	+	+	-
9	+	+	-
13	+	+	-
16	+	+	-
17	+	+	-
18	+	+	-
27	+	+	-
29	+	+	-
30	+	+	-
32	+	+	-
33	+	+	-
34	+	+	-
48	+	+	-
50	+	+	-
52	+	+	-
54	+	+	-
55	+	+	-
69	+	+	-
71	+	+	-
75	+	+	-
76	+	+	-
67	+	-	⊕
6	-	+	-
38	-	+	-
47	-	+	-
61	-	+	-
73	-	+	-
74	-	+	-
35	-	+	-
8	+	-	-
31	+	-	-
60	+	-	-
3	+	-	-
40	+	-	-
44	+	-	-
4	+	-	-
21	-	-	⊕
57	-	-	⊕
58	-	-	⊕

59	-	-	⊕
15	-	-	⊕
1	-	-	-
7	-	-	-
11	-	-	-
12	-	-	-
14	-	-	-
19	-	-	-
20	-	-	-
22	-	-	-
23	-	-	-
24	-	-	-
25	-	-	-
26	-	-	-
28	-	-	-
36	-	-	-
37	-	-	-
39	-	-	-
41	-	-	-
42	-	-	-
43	-	-	-
45	-	-	-
46	-	-	-
51	-	-	-
53	-	-	-
62	-	-	-
63	-	-	-
64	-	-	-
65	-	-	-
72	-	-	-
Correct	36	35	12 = 83
Errors	12	1	6 = 19

Coef. Rep. = 0.92

Coef. Scal. = 0.88

These results provide further confirmation of the orderings established earlier by Guttman Scale or Scheffe procedure. These orderings indicate that *ask* and *tell* constructions were relatively easy to interpret, although they were in a variable or unpredictable order (e.g. they had equal cell totals in the Scheffe test reported in Table 6.5 above). *Promise* sentences, on the other hand, were more difficult across all three levels of training. Thus, the respondents' performance could be described as highly systematic but variable; and their most favoured ordering can be summarized as follows:

**Tell > Ask > Promise.**

### **6.2.2. Implicational Scales for Universal Topicality Hierarchies**

The next step is to examine the patterns of behaviour among individuals at different levels, as regards the construct of 'topicality hierarchy' as referred to earlier. In Tables O, P, and Q (see Appendix I) these results are given. Since there appeared to be greater variability in the respondents performance, the criterion for accuracy was raised at 100 %. Thus, there were many violations of the main implicational patterns (as delineated in the tables), in turn, this led to lower or insignificant statistical results (i.e. for level 1, coef. rep. = 0. 81, coef. scal. = 0.71; for level 2, coef. rep. = 0. 77, coef. scal. 0. 63; for level 3, coef. rep. = 0.81, coef. scal. = 0.74). These results suggest that it would not be possible to establish particular orders of difficulty in relation to these combinations of features (except perhaps that both extremes of the range of combinations of features we could tentatively say which features were likely to be found). Generally, the easiest combinations comprise NPs of equal status on the topicality hierarchy, as in e.g. 'animate subject' combined with 'animate object' or 'human subject' combined with 'human object'. In addition, the most difficult combination is 'human subject' combined with 'animate object' or 'animate subject' with 'inanimate object'. These results also show that the pupils' performance differed markedly from that of both the university students and those at the teacher training college. As far as the pupils are concerned, the easiest combination is 'animate subject' with 'inanimate object' and the most difficult combination turns out to be 'animate subject' with 'animate object'. This was in fact the reverse of what both level 2 and 3 demonstrated in their performance. In order to find out more about this conflicting picture let us look at the actual sentences involved in the data. These were:

- 1.) The chicken asked the cat to come in. (animate x animate)
- 2.) The dog told the lion to wait. (animate x animate)
- 3.) The dog promised the cat to leave. (animate x animate)

These examples are precisely of the type in which both syntactic form and semantic information converge since both NPs have equal status. Therefore, it is not clear why these relatively less complex sentences should have been misunderstood, mainly by the least experienced respondents. Presumably, certain factors other than syntactic or semantic (but perhaps pragmatic ones) were brought to bear. In particular, as regards *ask*, it is possible to interpret *ask* as a request made by the chicken for the cat to come in (this would lead to a correct interpretation); or *ask* as an act of asking for the cat's permission to let the chicken in. In addition, in the pupils' educational environment, asking for permission is more common than making requests (i.e. from the pupils' viewpoint) whereas at the university or college, students are regarded as grown-ups and, thus, they are encouraged to take initiatives in relation to a wide range of domains. This fact may have lead the pupils to interpret some of these ambiguous *ask* sentences as if they involved seeking permission rather than a requesting to do certain things. In order to gain further understanding of some of these puzzling findings, we decided to analyse the type of errors of interpretation in connection with each verb type, levels and combination of subjects and object noun phrases. To this we turn in the next section.

### **6.3. Learner Language Behaviour**

In this section, we are going to look at both the correct and incorrect responses to each sentence type. Our objective is to examine the hypothesis that the status of both the subject and object NPs (in terms of the universal topicality hierarchy) affects the interpretation of the infinitival complements. It

has been established cross-linguistically that not all elements are equally likely to be selected as topics (Givon, 1979, 1984; Duranti 1979; Hawkinson and Hyam 1974; Morelong and Hyman 1977; Gass 1984). As we have already pointed out, the universal ordering of elements is as follows:

human > animate > inanimate

This implicational hierarchy means that there is greater likelihood for a human noun to be selected as a topic than for an inanimate noun to be selected as a topic (i.e. subject of the main clause). Gass (1984) proposes that this ordering is independent of syntax. Therefore it is a matter of theoretical and empirical import to find out precisely how the semantic information implicit in this topicality hierarchy interacts with language specific facts, in actual language learner performance. Recall that the Minimal Distance Principle (MDP) states 'the implicit subject of the complement verb is the NP most closely preceding it' (Chomsky, 1969: 10); and that this principle characterises the main pattern in English. It is therefore clear that one of the most crucial issues at stake in an investigation of the infinitival complements is to describe how potential semantic vs. syntactic conflicts are resolved when the elements of topicality hierarchy interact with certain applications of the MDP. From the learner language viewpoint, a dynamic dimension is necessary to pin down this resolution. We are going to examine the various sentence types at three different levels of training in order to capture the dynamics of the resolution of this potential conflict over time. For the sake of clarity and consistency in the presentation of the results given in the appropriate tables and figures, we shall look first at the number of responses to the first NP vs. second NP, across three levels of training. Secondly, we will provide the statistical significance of the results. Thirdly, we will summarize the findings in a graphical mode of

display, to show the impact of both levels of training and the role relationships of the two NPs involved (with particular reference to equal status, convergence or conflict of syntactic vs. semantic information). Additionally, and again for the sake of clarity, comparison of these verbs results with easier findings will be delayed until we undertake the discussion in section 6.4.

### 6.3.1. *Tell* Sentences

First we look at the three *tell* sentences in which NP<sub>1</sub> and NP<sub>2</sub> form and function are of equal status. This implies that both NPs are either 1) human/human, 2) animate/animate, or 3) inanimate/inanimate. Some of these results are given in Appendix Table R. The results are listed in terms of the correct response to the first NP (i.e. incorrect choice) vs. the second NP (i.e. correct choice).

Thus the figures 17-49 in the table for instance are to be interpreted in such a way that given the sentences,

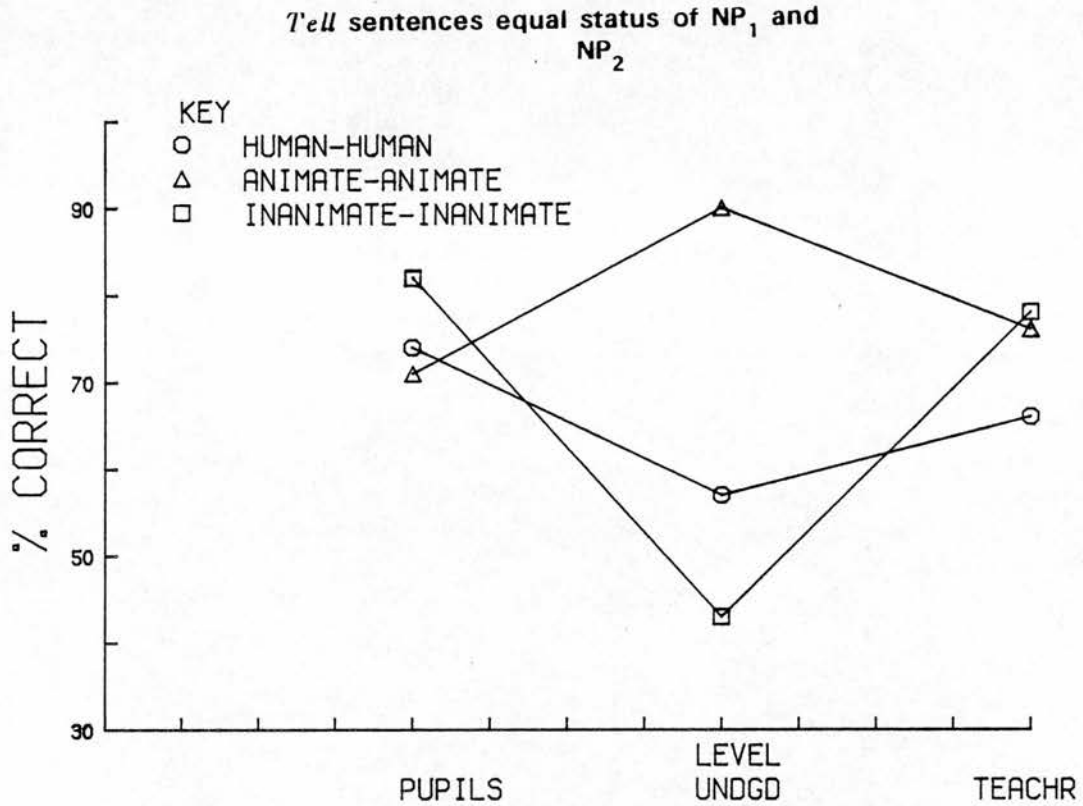
- 1.) The boy told the girl to leave.
- 2.) The dog told the lion to wait.
- 3.) The pen told the book to stop.

17 subjects at level 1 (i.e. 26%) responded incorrectly, that is, it is the boy who should leave, and 49 subjects (i.e. 74%) responded correctly, that the girl should leave. In addition, subjects at level 1 performed better than those at level 2 (i.e. 43 % error vs. 57 % correct) and level 3 (i.e. 34 % vs 66 %). The pattern for animate / animates shows that subjects at level 2 interpreted more accurately (90 %) the sentences with equal status NPs than the subjects at level 1 (71 %) or level 3 (73 %). These results are significant by Chi-square ( $\chi^2$



= 13.465;  $df = 6$ ;  $p < .036$ ). Figure 6.2 summarizes, in a graphical mode, these findings.

Figure 6.2

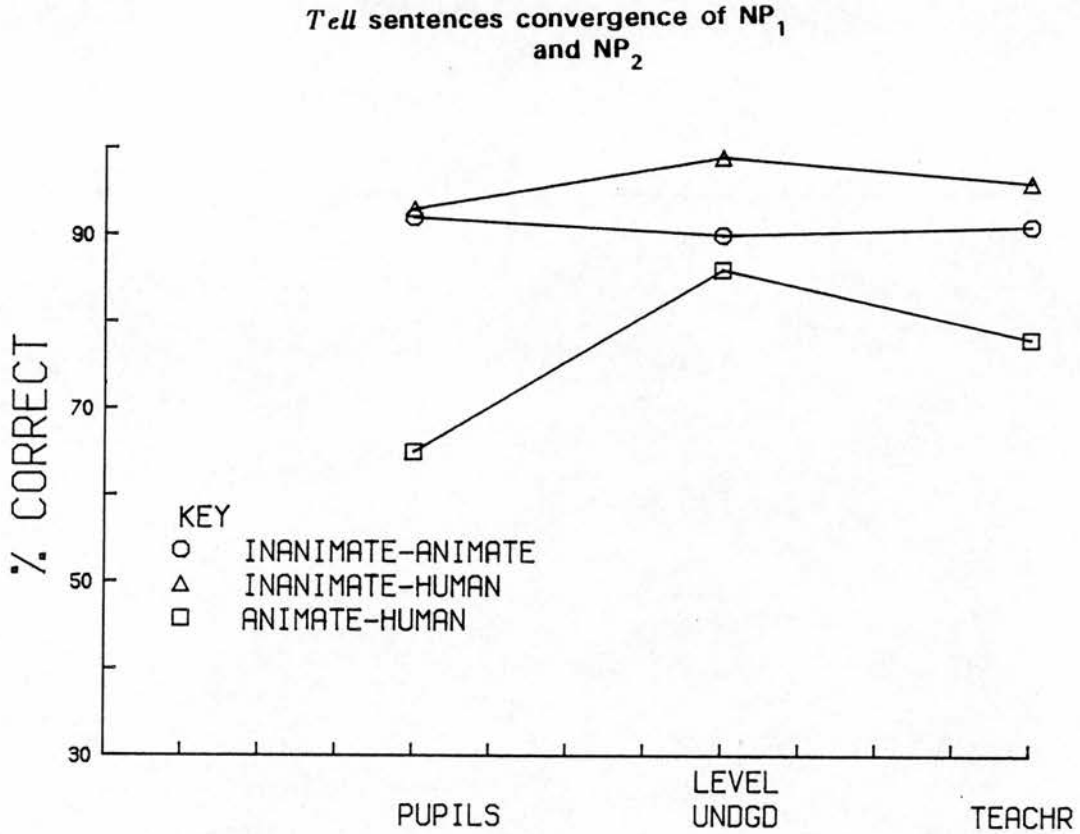


This figure suggests that from the perspective of both syntax and semantics, learners at a lower level of training did choose the correct NP (i.e. the second NP) as the subject of the infinitival complement more frequently than the subjects at level 2 and 3.

Secondly, we look at sentences involving convergence of form and function. In Table S (see Appendix I) are the results of those *tell* sentences in which the second NP is higher on the topicality hierarchy than the first NP; thus there is convergence of both syntax (in terms of MDP) and semantics. In a case like this, the prediction is that the subjects will select overwhelmingly the second NP as the doer of the action (in the complement clause). These results seem to

agree with the above prediction. However, there is no significant difference between the levels of training, suggesting that these sentences were quite easy for everyone ( $X^2 = 6$ ;  $df = 6$ ; ns.). These findings are summarized in Figure 6.3.

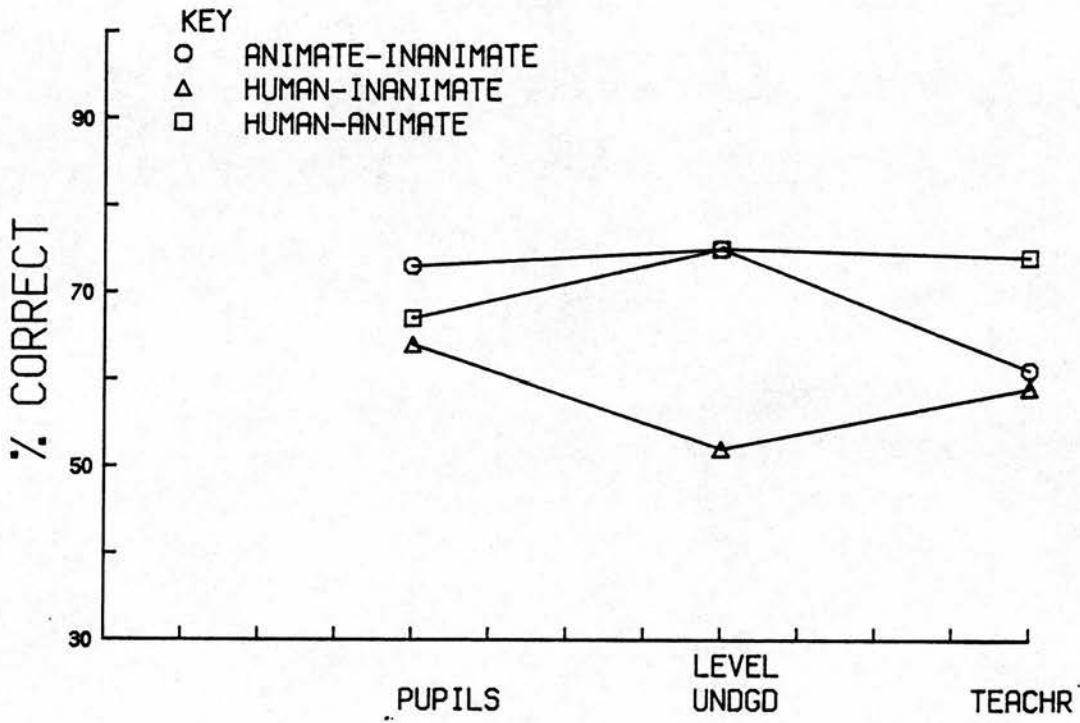
Figure 6.3



Thirdly we examined *tell* sentences which involve a conflict between form and function. These results are shown in Table T (see appendix). In these sentences NP<sub>1</sub> is higher on the topicality hierarchy than NP<sub>2</sub>; thus we can predict that these sentences will cause problems of interpretation. In other words, the most advanced students will conform to the general pattern of English, whereas the less advanced ones will tend to adopt a semantic-based interpretation.

Figure 6.4

*Tell sentences conflict between NP<sub>1</sub> and NP<sub>2</sub>*



The overall pattern across levels shows that the percentage of accuracy was much lower on this sentence type than when the two NP<sub>2</sub> have equal status or converging information. These percentages are 68 % for level 1, 68 % for level 2, and even less, 64 % for level 3. The Chi-Square results show that the levels did not perform differently in spite of the apparently poorer performance at level 3 ( $X^2 = 4.7$ ;  $df = 6$ ; n.s.). In addition, these findings are summarized in Figure 6.4 and Figure A (see appendix). Table U (see appendix) shows the relationship between level of training and overall performance (% correct) vis a vis the topicality hierarchy of NP<sub>1</sub> and NP<sub>2</sub> (i.e. convergence, equal status and conflict, in this order of increasing difficulty).

### 6.3.2. Ask Sentences

As with the *tell* verbs, we shall start first with those sentences involving equal status of form and function. This means that, given the sentences,

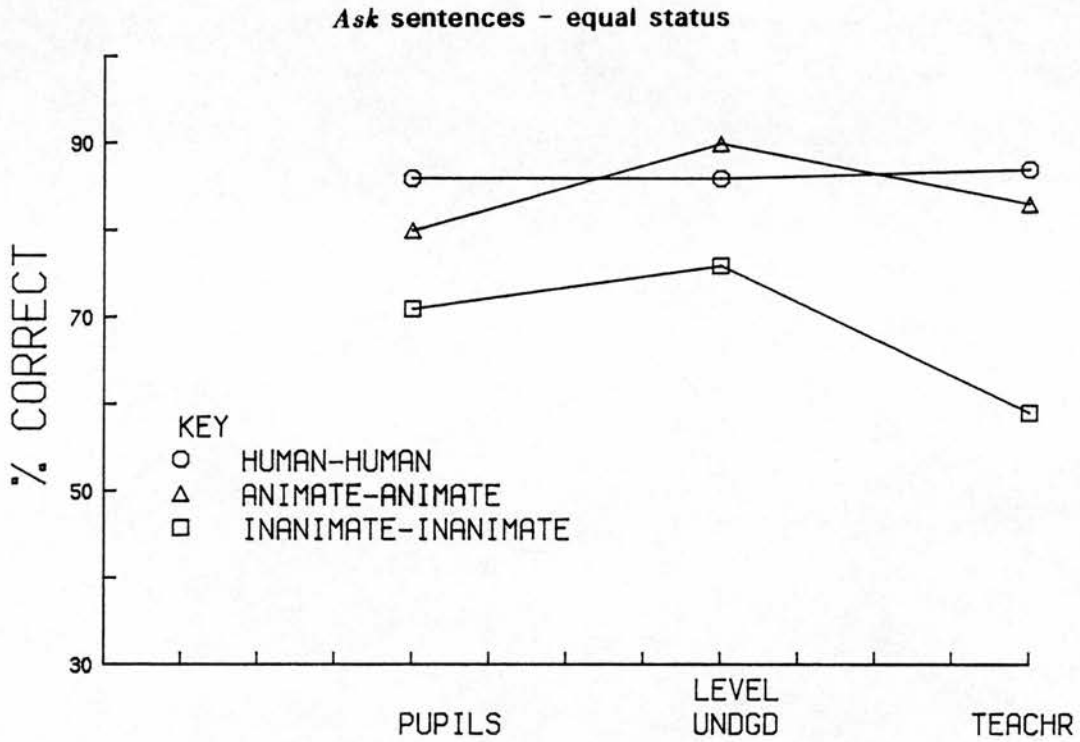
- 4.) The teacher asked the student to begin the lesson.
- 5.) The chicken asked the cat to come in.
- 6.) The cake asked the table to listen carefully.

at level 1, nine subjects (i.e. 14 %) responded incorrectly that (in e.g. 4) it is the teacher who should begin the lesson, and 57 subjects (i.e. 86 %) responded correctly that it is the student who should begin the lesson. These results are given in Table U (see appendix I). They suggest that overall, all three levels found the *ask* sentences fairly easy to comprehend; i.e. 79 % for level 1, 84 % for level 2 and 77 % for level 3.

In addition, these results suggest that in statistical terms, there was no difference between the way in which the respondents were performing at the three different levels of training ( $X^2 = 1.77$ ;  $df = 6$ ; n.s.). These findings are

summarized on Figure 6.5.

Figure 6.5

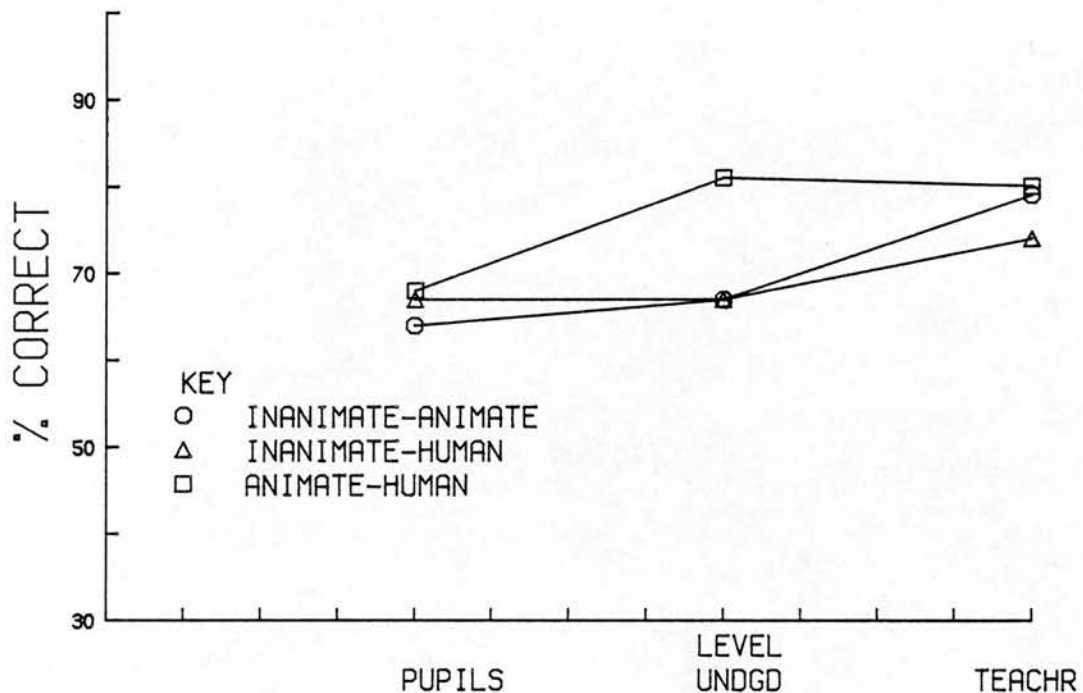


According to these findings, inanimate / inanimate sentence type proved more difficult; presumably since this set of NPs is semantically rather opaque (i.e. it is highly unlikely for inanimates to make requests), or because of the ambiguity of the *ask* sentences (i.e. *ask* can be understood as a request or seeking permission). Secondly, we examined the *ask* sentences in which form and function are converging. This is the case when the second NP is higher on the topicality hierarchy, thus in such instances, most subjects would have no difficulty in interpreting these sentences, provided that *ask* is understood as a request. These results are given in Table V (see appendix I).

These results seem to agree with the MDP which implies that NP<sub>2</sub> is the doer of the action requested by NP<sub>1</sub>. The more advanced students (level 3)

were making the correct choices of subject more frequently than the other two levels (i.e 78 % for the teacher trainees, compared with 71 % for the university students, and only 66 % for the pupils). These differences however did not reach statistical significance ( $X^2 = 7.8$ ;  $df = 6$ ; n.s.). Figure 6.6 summarizes the overall trend for *ask* sentences with converging syntactic and semantic information.

Figure 6.6

*Ask* sentences - convergence

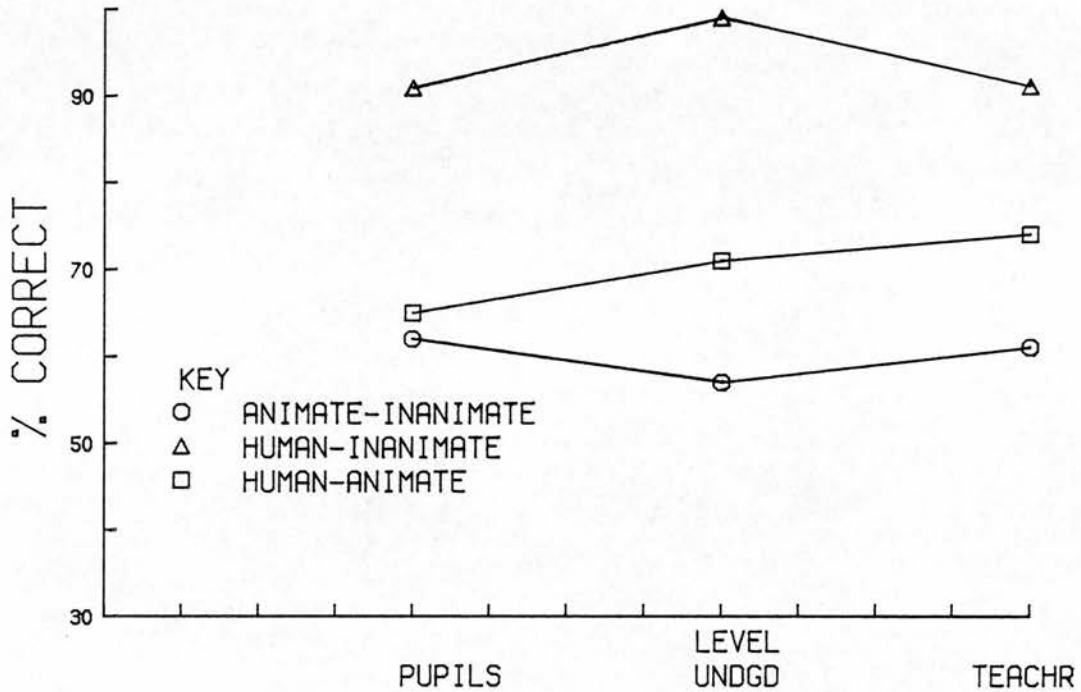
Thirdly, we look at those *ask* sentences in which form and function are in conflict. This implies the fact that the first NP is higher on the topicality hierarchy than the second. The resolution of such a conflict between syntax and semantics is hard to predict since *ask* itself has a somewhat ambiguous interpretation. Presumably, some of the more advanced subjects would favour a syntactic interpretation, whereas others (and especially the least advanced ones) would rely on the meaning. These results are given in Table W (see

appendix I).

Indeed these results show that the MDP (i.e. a syntactic interpretation) was preferred overwhelmingly by nearly all the respondents. This pattern is most clearly shown by the higher percentage of positive responses to the human / inanimate sentence type. These results are as follows: 91 % for the pupils, 100 % for the undergrads and 91 % for the teacher trainees. In order to examine the variable role of the topicality hierarchy, we look back at Table V which gives the results for the role relationships between NP<sub>1</sub> and NP<sub>2</sub>. The findings (in Table V, appendix) for the inanimate / human sentence type are as follows: 67 % for the pupils, 67 % for the undergrads and 74 % for the teacher trainees. So again the MDP is the 'winner'; and indeed, as Table W indicates, this time the winner takes all.

This trend is also implied by the Chi Square results since there was no difference between the three levels of training ( $X^2 = 4.03$ ;  $df = 6$ ; n.s). The results for the *Ask* sentences with conflicting information are summarized in Figures 6.7 and Figure B (see appendix I).

Figure 6.7

*Ask sentences - conflict*

The results in Figure 6.7 show that in the event of conflicting information between NP<sub>1</sub> and NP<sub>2</sub>, the human / inanimate sentences were the easiest (the actual sentence used was *The girl asked the ice-cream to eat quickly*); whereas the animate / inanimate type was the most difficult (the actual sentence was *The dog asked the door to come in*). The human / animate sentence type involved only average difficulty (as in *the visitor asked the pigeon to come in*). The fact that with the *ask* sentences the syntactic interpretation seemed to prevail suggests that, for these students, the meaning of *ask* is conceived of as a request, not seeking permission to do something. In other words if *ask* was being interpreted as seeking permission, then, we would find that a great proportion of the respondents selected the first NP as subject of the complement clause rather than the second. Instead we notice that, for human / inanimate, the



respondents prefer the second NP, which involves the interpretation of *ask* as a request. This preference of the MDP is shown unequivocally in Table W, in relation to human / inanimate sentences (i.e. 91 % for the pupils, 100 % for undergrads and 91 % for teacher trainees).

### 6.3.3. *Promise Sentences*

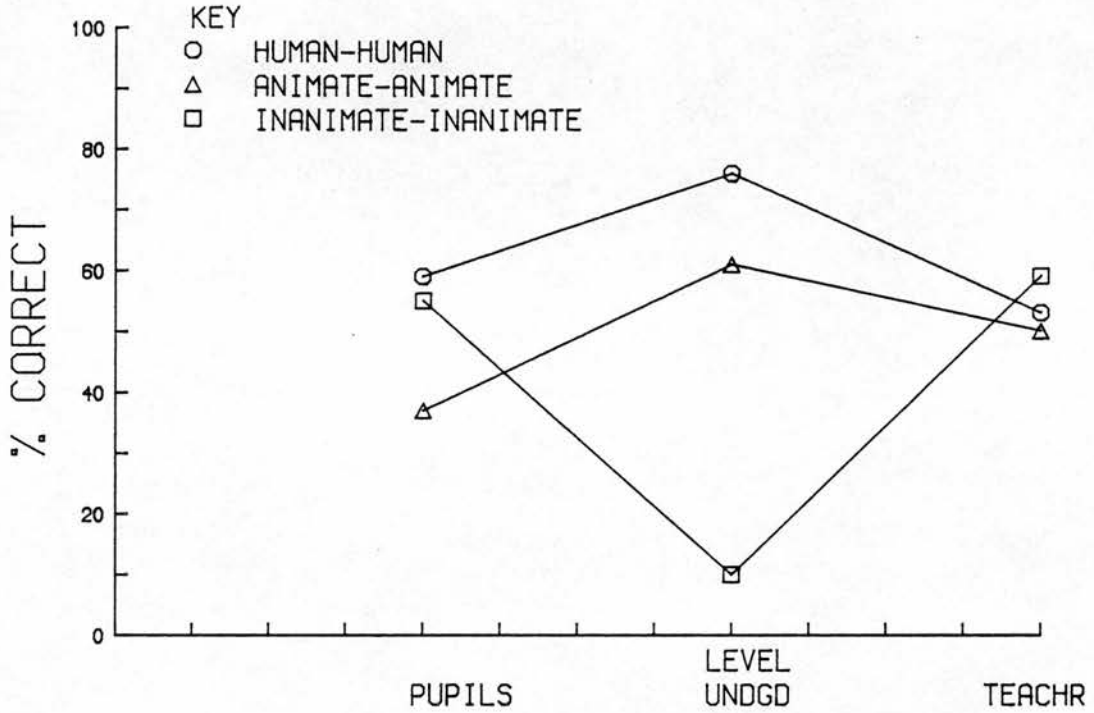
First we look at those sentences in which form and function have equal status. Sentences in this category include, for example,

- 7.) John promised Bill to study hard.
- 8.) The dog promised the cat to leave.
- 9.) The flower promised the tree to stay.

According to the results in Table X, the respondents are split 50 / 50 between the application and non-application of the MDP. Some 52 % of the total believe that it is John who should study hard, whereas the other half (i.e 48 %) think Bill should. However, there is a stronger tendency to break away from the MDP, and thus, to consider NP<sub>1</sub> as the subject for both the main and complement clauses. This moderate trend is shown by the Chi Square results ( $\chi^2 = 11.6$ ;  $df = 6$ ;  $p < .07$ ) which means that there is a tendency for the advanced learners to make a correct choice more frequently than the less advanced learners.

These findings are summarized in Figure 6.8. They suggest that the behaviour of the undergrads (i.e. level 2) was rather unpredictable, and thus, they may have been faced with greater difficulty.

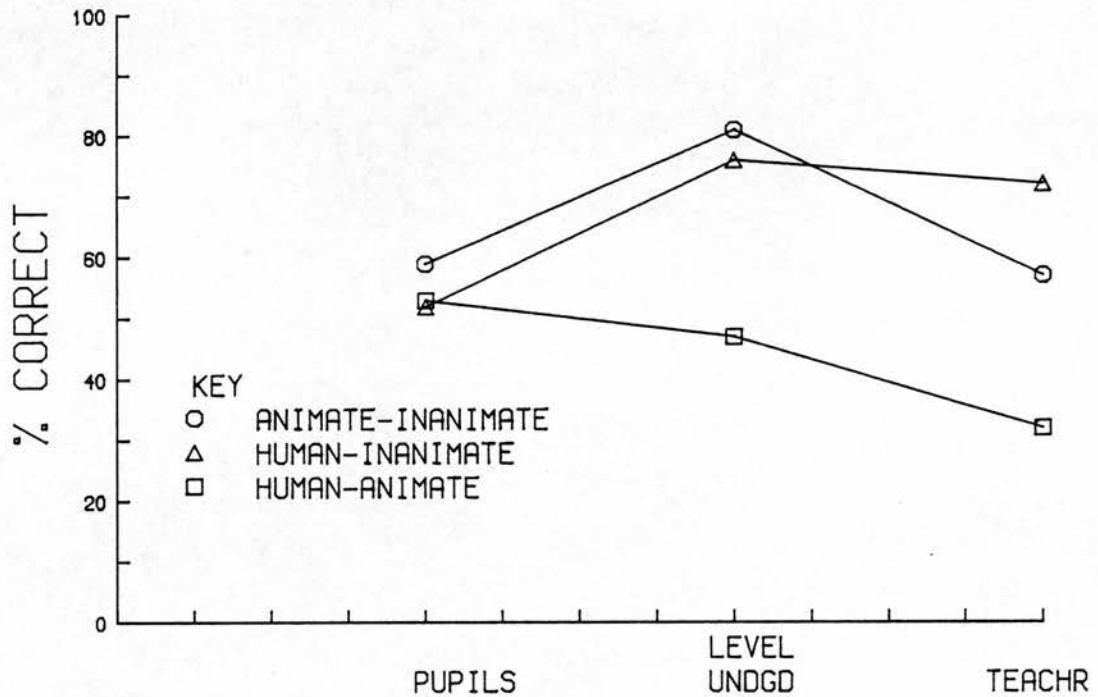
Figure 6.8

*Promise sentences - equal status*

Secondly, in Table Y, we examine the sentences in which there is convergence of form and function. In this case, the first NP is higher on the topicality hierarchy. Recall that *promise* verbs typically violate the MDP; and that choosing the first NP as subject would lead to a correct response, whereas with other verb types this would lead to an erroneous interpretation.

These results show that indeed there is a tendency for the respondents at all levels to choose the first NP as the subject of both the main and subordinate clauses ( $\chi^2 = 9$ ;  $df = 6$ ; n.s.); this implies that there were no differences due to the level of training. These findings are presented graphically in Figure 6.9.

Figure 6.9

*Promise sentences - convergence*

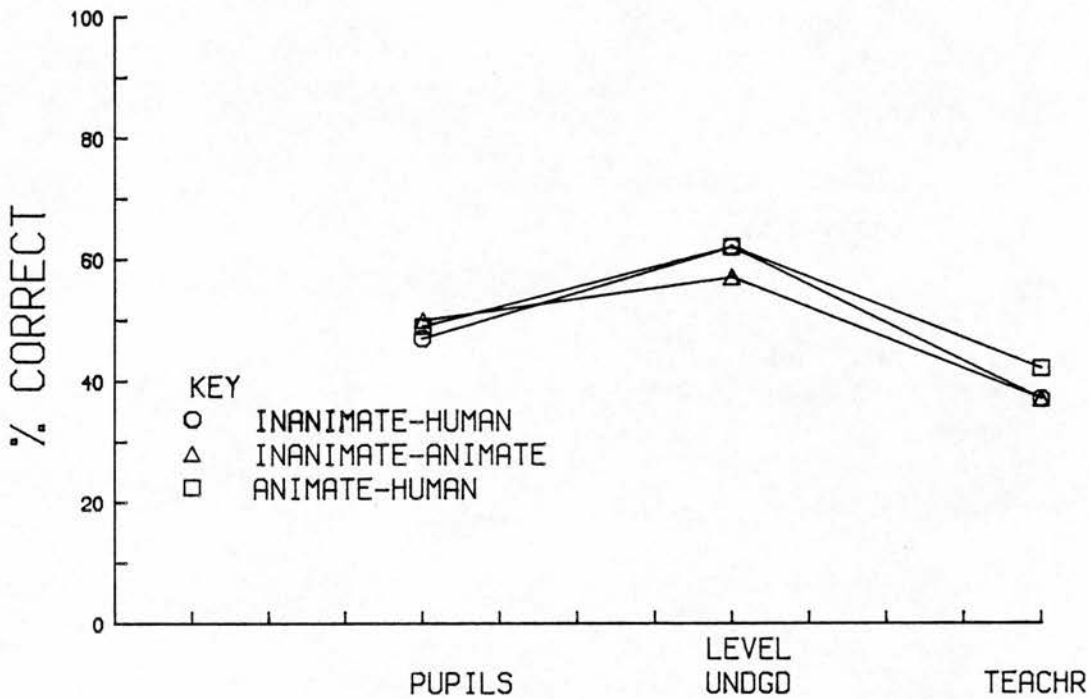
Finally we examine the *promise* sentences in which form and function are in conflict. This means that the second NP is higher on the topicality hierarchy. Selecting NPs as the subject of the subordinate clause would mean that the respondents are conforming to the MDP. With *promise* sentences, however, this choice would lead to an erroneous response. There is, then, a wide conflict between syntax and semantics, the resolution of which may require the violation of the established MDP pattern. These results are given in Table Z (see appendix I).

According to these findings, when there is conflicting information involving NP<sub>1</sub> and NP<sub>2</sub>, the pupils were totally confused; 51 % chose NP<sub>2</sub> as subject of the complement clause (which is incorrect, but in agreement with the MDP), and 49 % chose NP<sub>1</sub>. The university students, however, relied less on the MDP

(i.e. 40 % errors). The teacher trainees did rely considerably on the MDP, which led to the greater proportion of errors (61 %) at this level. These results, however, do not reach statistical significance ( $\chi^2 = 9.55$ ;  $df = 6$ ; n.s.) implying that at all three levels, the students are rather confused when it comes to choosing the correct NP subject of the complement clause. These results are summarized in Figures 6.10 and Figure C (see the latter in the appendix).

Figure 6.10

*Promise sentences - conflict*



#### 6.4. Discussion

Earlier studies of the so-called advanced structures have focused on the role of syntax, in terms of the linguistic complexity, in the comprehension of certain constructions (Chomsky 1969; D'Anglejan and Tucker 1975; Cooper, Olshtain, Tucker and Waterbury 1979). Our results provide some corroborative evidence to support the idea that syntax does play an important role in the

comprehension of these constructions. In addition, these results go even further to demonstrate the effects of the universal topicality hierarchies on the interpretation of the *tell / ask / promise* sentences.

The main body of our results are in agreement with those of Gass (1984) who found that, as regards the *tell* sentences, when there is no conflict between syntax and semantics the general tendency is for the correct English interpretation to occur. On the other hand, when there is conflict, semantics has a stronger effect at earlier stages, but syntax tends to win at later stages of development. While agreeing with the role of the universal hierarchies at any level, however, our data show less strong evidence across levels of training. In Figure A for instance, we see that the *tell* sentences are much easier when there is convergence of syntactic and semantic information, whereas they seem to cause greater difficulty when there is conflicting role relationship between the first and second NP. However, unlike studies in which subjects were L<sub>1</sub> children (Chomsky 1969), we do not find a definite pattern whereby level of proficiency could be regarded as a major predictor for determining which NP is most likely to be chosen as subject of the complement sentence. This may be due to the fact that both the children's conceptual development and syntactic maturation (the L<sub>1</sub> informants are usually under the age of 5) are both taking place at the same time. As far as the adult L<sub>2</sub> learners are concerned, conceptual development is already complete. This implies that L<sub>2</sub> learners rely both on principles of universal grammar, which can be re-activated by their language acquisition device, as well as their linguistic knowledge of the L<sub>1</sub> and the target language. Thus, the resolution of the potential conflict between syntax (i.e. form) and semantics (i.e. function) will depend, not only on change over time, but also on many other factors including transfer, generalization, knowledge of the world, learner strategies, and instructional variables.

According to our results, the level of training reached a moderate statistical significance only when NP<sub>1</sub> and NP<sub>2</sub> were of equal status (see Table :  $\chi^2 = 13.465$ ;  $df = 6$ ;  $p < 0.36$ ). In this instance, both the pupils (i.e. level 1) and teacher trainees (i.e. level 3) show greater awareness of the syntactic principles involved (i.e. MDP); and indeed they perform better than the university undergrads (i.e. level 2) who seem to depend more on the NPs role relationship or meaning. This is in effect the main trend found by Gass (1984) too although her results for *tell* sentences did not reach statistical significance.

The results of the *ask* sentences show a more homogeneous picture than what we anticipated. For these students, the second NP was always regarded as the subject of the complement sentence. Even when there is conflicting information between NP<sub>1</sub> and NP<sub>2</sub> (i.e. NP<sub>1</sub> is higher on the topicality hierarchy), the respondents still prefer NP<sub>2</sub> as the subject. These findings suggest that these students do not assign to the *ask* verb an ambiguous interpretation, i.e. both as a 'seeking permission' and 'request'. Only the latter interpretation would lead to the choice of NP<sub>2</sub> as the subject of the complement sentence (see Table W and Figure 6.7). In addition, there are no differences between the levels of training, which suggests that these students generally gave the *ask* sentences predominantly a syntactic interpretation.

As regards the *promise* sentences, the results are mixed. When NP<sub>1</sub> and NP<sub>2</sub> have equal status, we see (in Table X) that there is a tendency for the more advanced students to rely more on syntax than meaning ( $\chi^2 = 11.6$ ;  $df = 6$ ;  $p < .07$ ). However, when there is conflicting information then all of the respondents seem to be confused. In Table Z, this conflict is obvious among the pupils (viz 50 / 50 for the inanimate / animate NPs), and to a greater extent among the teacher trainees as well (viz. 62 % errors vs. 38 % correct). Surprisingly, the

university students who have been performing poorly so far, are doing better than the more advanced ones. The reason for this could be that the undergrads (i.e. level 2) are not sufficiently aware yet of the syntactic rule at stake, as far as the infinitival complements are concerned. Consequently, their responses tend to be more often affected by the semantic role relationships between the NPs involved, rather than the Minimal Distance Principle.

The results as a whole highlight a number of issues of both theoretical and empirical interest for second language acquisition research. First of all, the results suggest that we should not attribute single strategies to L<sub>2</sub> learners, even when we are dealing with one specific learning problem. Secondly, these findings call for further investigation into ways in which universals of language affect developing ILs, in particular, further studies should try to establish the strength with which these universals operate not only at sentence level but also at text and discourse levels. Thirdly, in relation to linguistic theory, a number of studies have attempted to establish implicational tendencies among language universals (Comrie 1984, Gass and Ard 1980). But in doing so these researchers have assumed equal distancing among the elements involved in these hierarchies. However, so long as the hierarchies are based solely on markedness (e.g. not taking full account of aspects of connected discourse) they may be linguistically real (i.e. in relation to grammatical competence), and yet lack psychological reality (i.e. pragmatic competence). Strictly speaking, the scope of our study does not enable us to tackle directly the issue of establishing psychological reality. This is so because our data deals with a rather limited sample of sentences, and thus may represent a reduction of what real people might do in real time in spontaneous discourse. Nevertheless, using appropriate quantitative techniques (viz. ANOVA, Scheffe test and implicational scaling) we have been able to show both the exact quantitative distancing

between the three verb types, as well as the preferred ordering according to which our respondents tend to interpret them. This ordering is as follows:

tell > ask > promise

Most studies in the literature agree on the fact that *tell* sentences (and others which conform to the MDP) are easier than *promise* sentences which violate the MDP (Chomsky, 1969; Karmiloff-Smith 1986; D'Anglejan and Tucker, 1975; Bongaerts, 1983; Cooper, Olshtain, Zucker and Waterbury, 1979). However, there have been conflicting claims as regards the *ask* sentences, mainly due to ambiguity (Chomsky 1972; Gass 1984). In addition, another way of tackling the issue of psychological <sup>reality</sup> in L<sub>2</sub> data would be to examine how L<sub>2</sub> learners proceed to reach the mastery of a function comparable to that of the target language, i.e. through the use of various strategies and changing hypotheses about the function of the NPs involved in the form of complement constructions. Many of these strategies are highly personal. However, some of the strategies adopted by a particular learner will soon overlap with those of other IL speakers. Thus the issue whether the paths followed (in arriving at the discovery of form to function relationships) are universal paths inherent in all language learning may remain unresolved. It may well be the case that they are not. Even if they were, however, the number of possible paths to follow is in principle, vastly greater than the paths actually followed (Zobl 1984; Huebner 1983). Other learners in a particular learning situation (or the same learners at different stages) could indeed follow more or less circuitous routes towards the discovery of the function of the complex predicate constructions in Standard English.

One major hypothesis which seems to emerge from these results is that the L<sub>2</sub> learner knows that language is intrinsically syntactic. This, in turn often leads the learner to adopt particular strategies in dealing with a learning



problem. In some cases, he may assume that, for instance, regarding the infinitival complements, the second NP is always the subject of the subordinate clause, i.e. if this is also true for his L<sub>1</sub> (Ellis 1986). In connection with most sentence types used in the present work, the respondents would find the right interpretation by relying on French (i.e. the main source of transfer).

Sometimes, though, the learners are in doubt, partly because of their fragile linguistic knowledge or because they are faced with an ambiguous construction (e.g. *ask* verbs involving other meanings than a request). When faced with a dilemma of this kind, L<sub>2</sub> learners may not be able to decide which NP qualifies as the subject of the embedded clause, and thus, they might adopt various strategies. Some high risk-takers may base their strategy on the assumption that 'if my interpretation of this construction is not similar to L<sub>1</sub>, it is probably correct'. Such a strategy was frequently used by adult L<sub>2</sub> learners in a study of tense and aspect reported on by Zydattiss (1976), which often lead to overgeneralization. Other more cautious learners may reason like this: 'Since this construction does not sound appropriate in my L<sub>1</sub>, it probably would not make sense in the target language either'. Sometimes these strategies will lead to a correct response, sometimes they will result in an incorrect one. For instance, the correct interpretation of the *ask* sentences seems to have been arrived at through the L<sub>1</sub> transfer strategy; since, in French, the indirect object (or dative NP) is normally the doer of the action in the infinitival complement (see Table V). Because of this possible reliance on French, the learners have not contemplated the alternative interpretation of *ask* (i.e. as asking for permission). If they interpreted *ask* in this latter sense, then the first NP would be preferred to the second NP as the subject of the subordinate. In fact, there is no evidence that at any level of training, the students preferred NP<sub>1</sub> to NP<sub>2</sub>, which would have been the case if the first NP were understood as 'seeking

permission' from the patient (i.e. the second NP) to take action. Other strategies may have been resorted to, especially when there was a conflict between syntax and semantics; and we shall not go into further detail here, since in many cases these strategies may turn out to be rather idiosyncratic in nature. The important thing to bear in mind is that such strategies are not fixed but may change with increased meta-linguistic knowledge or under certain communicative constraints.

## CHAPTER 7

### ASSESSING GLOBAL L<sub>2</sub> LANGUAGE PROFICIENCY

#### 7.1. Introduction

The work reported in this chapter establishes an objective measure for assessing global second language proficiency through the analysis of written compositions. The research design employed is based on the established assumption that although individual learner variability is evident in second language learning, there exist general developmental progressions to which learners conform (Larsen-Freeman, 1983). Over the last few decades both L<sub>2</sub> researchers and teachers have been concerned with the discovery of a reliable and readily applicable measure for assessing the development of second language competence in the classroom. In order to reach this aim, these classroom practitioners have had to find answers to the following questions:

1. Which measures would be successful for discriminating among the writing/ speaking abilities of ESL learners?
2. Would these measures increase over time as a function of ESL instruction?
3. If an L<sub>2</sub> index was found, to what extent would it be immune to the L<sub>1</sub> influence?
4. From a pedagogic viewpoint, can learners be taught how to write more sophisticated texts than their linguistic abilities permit?

In this chapter we shall attempt to answer these important questions as well as characterise the current state of the IL of Zairean students under investigation.

First we shall present a brief literature review of global assessment of language development. Second, we shall evaluate a number of concepts that have been used by both L<sub>1</sub> and L<sub>2</sub> researchers in this connection. We shall then

propose our own notion of a Communicational Capability Index (CCI) as an improvement on such earlier proposals as the T-unit.

Thirdly, we shall briefly describe our experimental design and then present the results of our quantitative and qualitative analysis of the Learner English of the Zairean students under investigation. We shall conclude the chapter by the discussion of the results and summary.

#### **7.1.1. A Brief Review of Literature**

Hakuta (1975) conducted a longitudinal study of the acquisition of English by Uguisu, a five year old native Japanese speaker. He then went on to compare Uguisu's use of relative clauses with that of Marta, a five year old Spanish speaker learning English as a second language, who was then a subject of a study by Cancino, Rosansky and Schumann (1975). One of the issues Hakuta was interested in examining was language transfer as a factor in the second language learning process. More specifically, Hakuta intended to examine the phenomenon of structural avoidance observed by Schachter (1974). Hakuta hypothesised that since in Spanish, as in English and unlike Japanese, the relative clause occupies the position to the right of the head noun. Marta, the Spanish speaker, would produce more relative clauses than Uguisu, the Japanese speaker. He was unable to test this hypothesis however, since he had no reliable means of determining whether Uguisu and Marta were at a similar stage of development in English. Marta did produce more relative clauses than Uguisu. However, it would be quite plausible to account for this finding on grounds that Marta had a better command of English, and not necessarily because she spoke a first language more similar to English than Japanese. If it were possible to establish that both children were at the same stage of development in English, their second language, then it would be more

acceptable to advocate the use of a transfer-induced avoidance phenomenon or strategy on the behalf of Uguisu, the Japanese speaker.

Both Taylor (1975) and Larsen-Freeman (1975, 1976) have argued that the strategies of the learner or their influences on the learning process change as the learner's proficiency in the target language increases. Taylor (1975), for instance, administered a translation test to 20 Spanish-speaking subjects of English as a Second Language. The subjects were at two levels of proficiency, 'beginning' and 'intermediate'. An error analysis made of their translations revealed that the errors committed by the elementary and intermediate level students were not qualitatively different. However, their relative influence depended on the level of proficiency: the elementary subjects' reliance on the transfer strategy was found to be significantly higher than that of the intermediate subjects.

Larsen-Freeman and Strom (1978) examined 48 compositions written by non-native speakers in an attempt to delineate a second language index of development. The compositions were impressionistically allocated to five proficiency levels by two independent researchers in order to identify which features made each level unique. The features considered included writing mechanics, clarity, organization, grammar, lexical choice, number of words, number of T-units (Hunt, 1965) average length of T-units, number of error-free T-units, sentence construction and content. Larsen-Freeman et al. found that the measures which seemed most suitable as a basis for an index of development, were the average length of the T-unit and the total number of error-free T-units per composition. One of the most significant findings of the Larsen-Freeman and Strom's (1978) study was that the compositions they had evaluated as 'poor' exhibited fewer errors in article usage, for instance, than the

compositions they had evaluated as 'fair'. This finding seems to bring further support to the idea that language learning need not, and indeed does not progress in a linear way (Meisel, Clahsen and Pienemann, 1981). Neuman (1977) conducted a thorough error analysis of compositions in an attempt to characterize the problems of learners studying English as a Second Language at the intermediate level at UCLA. Neuman too concluded that errors cannot be used to distinguish the intermediate level from the beginning and advanced levels. In view of the findings documented by the above studies, Larsen-Freeman and Strom (1977) decided to turn to more objective measures of length. In the next section, we shall evaluate some of the established constructs often used to assess L<sub>2</sub> development.

## **7.2. Some Theoretical Considerations**

### **7.2.1. The Notion of Mean Utterance Length (Brown and Fraser, 1964)**

Child language acquisition research in the 60's manifested the same need to establish an objective measure whereby first language development could be accurately gauged. Brown and Fraser (1964: 72), for instance, used the Mean Utterance Length (MUL) to study the acquisition of thirteen grammatical morphemes by children acquiring English as their first language. Although Brown and Fraser (1964), like McCarthy (1954) earlier on, were able to find an age-related increase in the mean length of utterances, it has been argued that the MUL could not be applied to adults or older children (Huang 1970). As far as the children involved in the Brown and Fraser's (1964) study were concerned, age ranged between 26 and 36 months, and their MUL also increased gradually from 2.0 to 4.9 morphemes or words, in keeping with the children's age.

### 7.2.2. The Notion of T-Unit (Hunt, 1965)

Huang (1970) used the mean utterance length (MUL) in his study of the acquisition of negatives and questions by a Taiwanese child learning English. However, Huang found this instrument unsuitable for older children, and therefore for adults too, since this child could produce many long utterances right from the beginning stage, such as, e.g. *It is time to eat and drink*. Because of this fact, Larsen-Freeman and Strom (1977) preferred the notion of a T-unit to analyse their written composition data.

Originally devised by Hunt (1965), a T-unit is defined as a 'minimal terminal unit. . . minimal as to length, and each . . . grammatically capable of being terminated with a capital letter (sic) and a period.' (Hunt 1965, quoted by Larsen-Freeman et al., 1977: 128). According to this definition, more than one T-unit may occur within a single sentence. However, as the following will show, the number of T-units alone cannot enable us to assess the embedding process. For example,

- 1.) I first met this very special man when I was only about six years old, and my earliest memories of him are vague.  
(i.e. Two T-units, with total embedding weight of four in terms of Bever's (1972) perception theory.)
- 2.) I've got a dog and his name is Blue.  
(i.e. Two T-units, but with total embedding weight of two.)
- 3.) I've got a dog named Blue.  
(i.e. One T-unit; but its total embedding value is three.)

Thus, in their analysis of compositions, Larsen-Freeman and Strom (1977) attempted to establish first, the total number of words, the number of sentences and the number of T-units per composition. Then, they calculated the number of error free T-units and also the average word length per T-unit, since according to Hunt (1965), as children mature the number of sentences

they write diminishes but the length of the T-units increases due to embedding and other processes which demote main clauses to a subordinate status. So, by calculating the average word length per T-unit, Larsen-Freeman et al. believed that this would enable them to examine the subordination or embedding process that Hunt (1965) had in mind. However, it is quite conceivable to have two sentences with the same number of words, and yet with very different sentence embedding patterns, or even a short sentence with greater complexity than a longer one, as in (2) and (3) above. It is fairly obvious that the embedding process would be appropriately examined by focusing on the number and type of subordinate clauses within the sentence; not by counting the number of words. To this extent, the second language proficiency indices based upon a T-unit will only be as good or as bad as the analysis upon which these indices are based.

Another distinctive feature that Larsen-Freeman and Strom (1977) attempted to establish in order to delineate L<sub>2</sub> proficiency was composition length. The authors speculated that longer compositions (in words) might be due to a better command of syntax or vocabulary. However, they had also contemplated a more cautious interpretation of data by admitting that longer length of text in words might be attributed to the 'willingness of some subjects to be more expressive and, thus, take greater risks of committing errors.' An analysis of variance of their data, however, revealed non-significant results in relation to composition length.

The main difficulty with the notion of a T-unit in general, and its interpretation by Larsen-Freeman et al. in particular, was highlighted by one of their findings (1977: 130), namely that

A couple of 'excellent' compositions contained short T-units. One



of these had excellent grammar and spelling but not as much conjoining or relativizing as there could have been. The other used a great deal of dialogue, which made the composition very interesting but did not allow for long T-units.

This shows the kind of dilemma researchers have to contend with in deciding which criterion (i.e. accuracy or fluency) ought to be adhered to in judging running text or speech as 'bad' or 'excellent'. More importantly, as far as the present study is concerned, it would appear that the conceptualization of a T-unit itself leaves much to be desired, that is as a 'minimal terminable unit . . . as to length. . . and grammaticality'. From this standpoint, therefore, it might be in order to explore better analytical techniques such as those which have been proposed by researchers working in the area of information science, speech perception, and case grammar (Bever, 1972; Pepinsky, 1974; Chafe, 1971; Andersen, 1971; Cook, 1973). This will lead us to the illustration of these techniques using the case grammar matrix model (Cook, 1979), and then, we will propose the notion of Communicational Capability Index (CCI) as a more accurate and satisfactory way of assessing second language proficiency.

### **7.2.3. Bever's (1972) Speech Perception Strategy**

Bever (1972: 104) reports on investigations into the structure of speech perception and concludes that 1.) the clause is the primary perceptual unit; 2.) within the clause, direct mapping rules assign semantic relations between major phrases; and 3.) that after each clause is processed, it is recoded into a relatively abstract form, thereby leaving immediate storage available for processing the next clause.

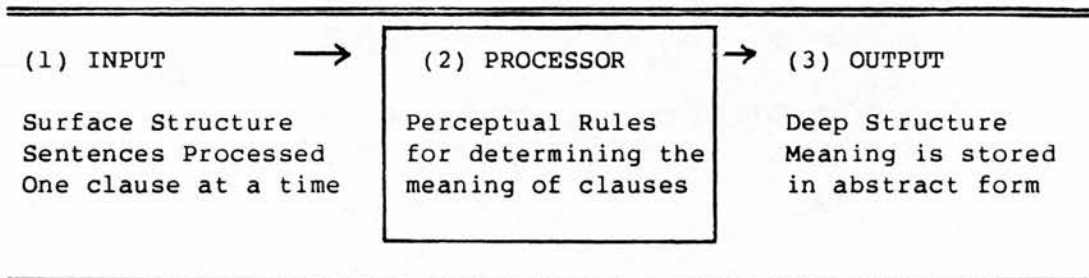
The theory that the clause is the primary unit of information was empirically tested by interrupting the flow of speech by simple clicks. In a series of experiments it was shown that: 1.) reaction to clicks is faster at clause

boundaries, 2.) clicks are accurately located when they occur at clause boundaries, and 3.) clicks are mislocated when they occur at other than clause boundaries. From these facts, Bever (1972: 105) concludes that

. . . during a clause we accumulate information and hypotheses concerning its deep structure; at the end of the clause we decide on the structure of what we have just heard.

The theory that clauses are erased from short-term memory after processing and stored in an abstract form was tested by asking subjects to process clauses in groups of two. The results of this experimentation were that 1.) recall of the meaning of both clauses is virtually perfect, but recall of the words of the first clause is worse than recall of the second clause; 2.) words from the second clause are identified faster than words from the first clause, and 3.) the structure of clauses is forgotten after a few clauses.

*Figure 7.1: Speech Perception (Bever, 1972)*



As Bever indicates in Figure 7.1, the input to the speech processor is the surface structure of the sentence, but sentences are processed one clause at a time. Within the processor, perceptual rules interrelate major phrases to determine the meaning of each clause. The surface structure of the clause is then erased from short-term memory and the meaning of the clause, i.e. its deep structure is stored in the memory in abstract form. The recognition of the clause as the unit of information in speech processing seems to warrant the suggestion that the clause (not individual words) should form the basis for the

analysis of particular texts and for labelling them as being more or less advanced or more or less difficult to process with regard to the number and kind of clauses present.

#### **7.2.4. The Notion of Information Blocks (Pepinsky, 1974)**

Pepinsky (1974: 59) reports on the findings of a research project entitled *A Metalanguage for Systematic Research on Human Communication vis Natural Language*. Using a multi-disciplinary team of researchers at the Ohio State University, Pepinsky was able to establish the clause as the unit of information. The findings of the research went further than that of Bever (1972) in establishing between clause and sentence an intermediate unit of information called the 'Information Block'. This block consists of a clause group clustered around a single main clause. Pepinsky could thus reduce texts by computer analysis to information display. All coordinating conjunctions between clauses would be treated as block boundaries. As a result, both sheer sentence length and the conjoining process used to join clauses could be eliminated as significant factors in this analysis.

Applied to Bever's perception theory, Pepinsky's proposal suggests that the speech processor processes not only single clauses, but also clause cluster centred around a single main clause. Thus, two coordinate clauses can be interpreted independently one from the other, whereas, for instance, subordinate clauses can only be interpreted in the light of the main clause on which they are dependent. A relative clause, for instance, which modifies a noun in the main clause must be processed before the antecedent in the main clause is erased from short term memory.

Compared to Hunt's (1965) T-unit, Pepinsky's (1974) Information Block Theory offers greater analytical precision, although prima facie they seem to be

capable of leading to the same outcomes. The crucial factor lies in the way according to which each of these two theories (i.e. information block theory and T-unit concept) enables us to delineate the different clause types. A closer look at both concepts, in the next section, will reveal that different interpretations of the notion of a clause could, in practice, affect the outcome of an analysis based on the principles of clause as a unit of information.

#### 7.2.5. The Notion of Communicational Capability Index (CCI)

Cook (1979) defines a clause as a 'string of words containing one and only one predicate'. Thus, there are as many clauses as there are verbs acting as predicates, and word groups clustered around an infinitive, participle or gerund are counted as clauses. However, depending on the context, verbless clauses may occur. In addition, simple clauses with more than one verb can also occur. In order to illustrate this interpretation of a clause, let us consider the following dialogue which took place between speakers A and B. A, who lives in Edinburgh, has just arrived in Spain to visit his friend who lives in a small fishing village there:

- (1) A: Who is the best known fisherman in the village?
- (2) B: Eduardo . . . the old man with a wrinkled face.
- (3) A: How many fish does he catch per day?
- (4) B: A lot. Some are as big as a cricket bat.
- (5) A: Yesterday, I saw a fisherman with deep wrinkles on his face. He was returning home with his bag empty.
- (6) B: Really?
- (7) A: Yeah. He looked thin, jaded and emaciated.
- (8) B: That's him. . . Eduardo is an old man who fishes alone and has gone 18 years without a holiday. Sad, isn't it?
- (9) A: Yes, it is.

Applied to the above (imaginary) dialogue between A and B, the principle of one verb – one clause fails to identify many phrases that we could still regard

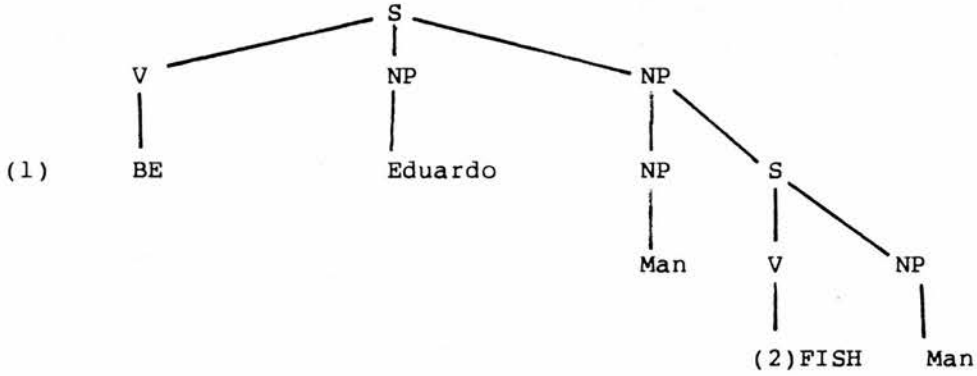
as clauses, although this principle does indeed apply in most cases. First of all, this principle cannot be applied strictly in cases where two or more verbs may form a single clause as in (7), where two verbs (i.e. *looked* and *emaciated*) and two adjectives (i.e. *thin* and *jaded*) are conjoined to form just one single predicate. Secondly, the principle of 'one verb, one clause' would fail completely to account for verbless clauses, e.g. minor sentences with no verb, short replies as answers to questions, etc. (as in (4) and (6)). Moreover, there are verbless clauses where the verb must be retrieved or supplied from the syntax or deep structure. These verbless clauses include: 1) comparatives (as in (4) *Some are as big as a cricket bat*); 2) manner phrases introduced by *with* or *without*, or which contain a postponed adjective (as in (5) . . . *with his bag empty* or *he has gone 18 years without a holiday*); and 3) locative phrases (as in (5) *I saw a fisherman with deep wrinkles on his face*).

However, if this broad interpretation of a clause (similar to that of Cook (1979)) is combined with Pepinsky's (1974) theory of information blocks, then, we can measure accurately the clause patterns characterizing written texts, in a better way than the theory of a T-unit can do. The reason for this is that the T-unit seems to depend too much, and almost solely, on the surface structure of the sentence and seems to pay little attention to its deep structure. For instance, if we take sentence (2) above (*Eduardo . . . the old man with a wrinkled face*), it is not clear whether such an utterance could qualify as a T-unit, that is, a terminable unit in a grammatical sense, or even a clause. And yet, we would argue that the deep structure sentence of (2) does include two information blocks, or two independent clauses (viz *Eduardo is an old man* and *Eduardo has a wrinkled face*). Using the insights gained from Cook's (1979) case grammar matrix model, it can indeed be shown how such clauses do form a normal information unit in the deep structure. Let us take example (8) above:

*Eduardo is an old man who fishes alone and has gone 18 years without a holiday.*

*Figure 7.2: An Illustration of the Informative Display in Deep Structure*

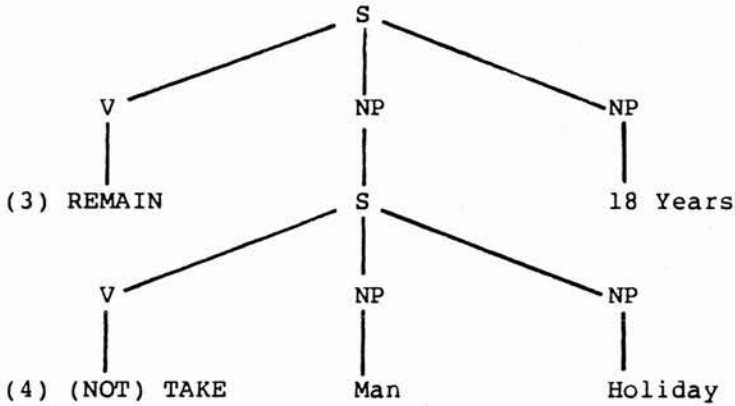
Block 1:



*Eduardo was an old man who fishes alone*

+ And

Block 2:



*Eduardo has gone 18 years without a holiday*

As shown in Figure 7.2., the information contained in sentence (8) is displayed as two 'blocks' comprising two clauses each. In the first block, there is one main clause and one relative clause. In the second block, there is one

main clause and a manner subordinate clause. The relative clause in block one is understood in the context of the main clause. The manner clause, in turn, is understood in the context of the block containing the third and fourth clause. We would also add that the fact that the two blocks are joined in a single sentence is immaterial; so too is the number of words involved. Following a suggestion made by Cook (op. cit.) we would also argue that the complexity of any written text would depend primarily on the number and kind of clause embeddings (i.e. the embedding weight of its information blocks). To this extent, the breakdown of a text into clauses can be regarded as the reverse of the embedding and conjoining processes by which clauses were combined into sentences in the first place.

According to Transformational theory, a sentence which contains embedded clauses is processed one at a time, beginning with the lowest embedded clause, then the next higher clause, until the main clause is reached. This process of moving from the lowest to highest clause has been referred to in Chomsky's Extended Standard theory as the 'bottom-to-top' principle. Formally speaking, this is a cyclic principle whereby

Any rule-application whose domain is D must precede any rule-application whose domain of application includes (= dominates) D.

Radford, 1981: 201

Both case and transformational grammars have enabled us to propose the notion of a Communicational Capability Index (CCI) which can be defined as an objective measure of the ability of a particular learner to encode messages or process information in his target language. The extent to which the learner actually uses this ability cannot be determined by CCI; since it is intended only as an index of the elaboratedness that the learner is capable of, not how he

makes use of his CCI, whether successfully or not.

In analytical terms, this notion of CCI can reveal a great deal about a text's communicational ability (i.e. its conceptual content), since it focuses on the way in which a second language learner combines words and linguistic units together in an attempt to create, ultimately, a coherent text. Of course linguistic units do not have values in isolation. They assume particular values through their relationships with co-text and situational context to form coherent discourse. The establishment of discourse values through the relationships between linguistic units in context has been at the core of many specialists' debates including both Speech Act theorists (Austin, 1962; Searle, 1969, 1979; Levinson, 1983) and those working in the ethnography of speaking (Sacks, 1972; Bauman and Sherzer, 1974; Sacks, Schegloff and Jefferson, 1974). One of the facts that has emerged from these studies has given strong support to Wittgenstein's (1958: 10-11) view, viz. that 'there are as many discourse values (or speech acts) as there are roles in the indefinite variety of language games (or speech events) that human beings are capable of inventing'. This view highlights that the analysis of conversational interaction or of various genres of texts is bound to be not only a fascinating enterprise but also a very complex one. In my view, however, those approaches which have given adequate account of language as communication, or coherent discourse, are more promising than those which limit themselves to morphological and syntactic considerations (e.g. Widdowson, 1979; Crombie 1985a, b; Longacre 1972, 1976).

According to Crombie (1985b), sentence meaning is determined largely with respect to the interaction between words and structures in which they occur; discourse value is determined largely with respect to the interaction between



sentence meaning (i.e. conceptual content) and context. Discourse values may be divided into two types: unitary values (e.g. request, warning, threat, insult, etc.), and binary values (e.g. reason-result; purpose-means; condition-consequence, etc.). According to Longacre (1972: 52), there appear to be a limited analytically manageable number of binary values and these have a high degree of comparability across languages. Additionally, since binary values are very often lexically and syntactically signalled, the study of binary values has direct implications for the study of vocabulary and syntax.

Clearly, there is a good deal that could be said or done about the analysis of text or written composition but which will not be said or done through the use of our proposed notion or technique of assessing the communicational capability index among  $L_2$  learners. The main reason for this is its limited scope. Another important reason is the criterion of economy. In this connection, we feel that if our attempts are to succeed in establishing the relationships between linguistic units in context, it is important to begin such a task from as solid a linguistic base as possible and at another stage, to concentrate on so-called discourse values.

Thus our own attempt to establish a Communicational Capability Index (CCI) is, in a sense, a first stage towards establishing semantic relations between propositions, i.e. a stage towards the analysis of truly coherent discourse. While we believe that this measure can be reliably applied to all  $L_2$  learners, we are also aware of the fact that a lot more could be done. However, as we pointed out above, only reasons of economy of time and space have imposed upon us the methodologically sound constraint of focusing only on one specific step of text analysis, with a view to find a global index for assessing proficiency in a second language. In addition, it seems to us that by focusing

on clauses (rather than words, as did Larsen-Freeman, 1983, and Larsen-Freeman et al., 1977) we would get a better idea of the learner's ability to encode messages which is just one, but an important stage in communication.

### 7.3. Design

In the preceding section, we have examined a number of ways in which both  $L_1$  and  $L_2$  researchers have attempted to establish a reliable measure for assessing global language proficiency. Various concepts were discussed including the Mean Utterance Length (Brown and Fraser, 1964), T-unit (Hunt, 1965), Speech Perception strategy (Bever, 1972), Information Block Theory (Pepinsky, 1974) as well as that of a Communicational Capability Index (CCI). The latter concept was proposed as a way of resolving some of the analytical inadequacies encountered in Hunt's T-unit, or at least its interpretation by Larsen-Freeman (1983) and Larsen-Freeman and Strom (1977). In this section we will present the different steps that ought to be undertaken by the experimenter in order to establish the Communicational Capability Index of a  $L_2$  learner through the analysis of his written composition.

If, as stated earlier (Bever, 1972, Pepinsky, 1974), clauses constitute information units and this information is communicated in blocks clustered around each independent clause, then, it is of crucial importance to find an analytical technique which can isolate clauses in running text. Our main design requirement will then be to identify the sentence patterns within the data, in terms of (1) its constituent clauses, (2) number of information blocks, (3) clause embedding weight, (4) error-free information blocks, and (5) finally, its CCI.

In order to obtain these different measures, the following steps should be followed. First of all, the text must be broken into single clauses and practically

rewritten one clause to a line. This is the reduction process (Cook 1979:43 ) and is in effect, the reverse of the embedding or coordinating processes. Secondly, the clauses thus identified within each sentence should be marked by a letter (A, B, C, D . . . ) so that the value of clause embedding in the text could be calculated at the appropriate stage. The main clauses will be marked (A) whereas the subordinate clauses will be marked (B, C, D, etc.). Thirdly, clusters containing one A-clause and B, C, D, . . . clauses constitute one information block. The symbol (+) is used to separate A-clause clusters within the sentence. The symbol (£) is used only at the beginning of a sentence. Thus, it is fairly easy to calculate the different measures that are necessary in order to find the index of communicational capability manifested in the learner's texts. These different measures can then be used in an analysis of variance and other procedures to test for significant differences between subjects across the different levels of training. The list of the different measurements or variables required for the analysis is as follows. We must calculate:

- 1-The total number of sentences within the text.
- 2-The total number of clauses within the text, both main and subordinate clauses included.
- 3-The total number of information blocks, i.e. the number of A-clause clusters per sentence.
- 4-The average number of clauses per sentence.
- 5-The average number of clauses per information block.
- 6-The total value of clause embedding weight. To represent embedding weight, numerical values are assigned to clauses, with A=1, B=2, C=3, D=4, etc. The total value for the text sample is, then calculated by multiplying the number of A-clauses by 1, the number of B-clauses by 2, the number of C-clauses by 3, and so on. In psychological terms, C must be held in memory for a D-clause if D depends on C, B must be held in memory for a B clause which depends on A. Thus, the D-clause takes, in theory, four times as long to process as the main clause A, C takes three times as long, and B twice as long (Cook, 1979).

- 7-The average clause embedding weight (ACEW), that is, the total value of embedding weight should be divided by the number of clauses in the sample (see (2) above). A text which contains only A clause would have an ACEW of 1.00. Any average above this figure indicates the additional processing time which is required on the average for each clause in the sample. For instance, an average of 1.88 means that 88% more time is required to process clauses in the sample than would be required in simpler texts.

Of course all embedded clauses would not conceivably take the same amount of processing time. However, since the ACEW is an average, it is hoped that absolute differences in processing time for different kinds of embedding (even if they were possible to predict) would cancel out and leave the average clause depth (or embedding weight) relatively unchanged.

These measures have been successfully used to assess the complexity of selected texts, ranging from simple ( e.g. Mark Twain's *Tom Sawyer*) to medium (e.g. Graham Greene's *I Spy*) and complex texts (e.g. Dostoevsky's *White Nights*). Moreover, Cook (1979) found that the average block length (see (5) above) correlated positively with the average clause embedding weight (ACEW) for each of the 36 writing samples selected for his investigation.

In addition, two more steps (i.e. 8 and 9) have to be undertaken before we can establish what we would regard as the L<sub>2</sub> learners Communicational Capability Index (CCI), both steps involving the notion of error-free. Following a suggestion made by Gaies (1976), we must acknowledge the fact that L<sub>2</sub> learners (unlike older L<sub>1</sub> learners of native speakers) regularly commit lexical, syntactic and morphological errors. Some of these errors affect communication but others may not; and if they do some could be more crucial than others. In order to circumvent this drawback, it is again more plausible to use proportions rather than absolute numbers of error-free clauses or clause-clusters. As Gaies (1976: 7) goes on 'some account must be taken of structural errors which occur in L<sub>2</sub> learner's writing, for they, too are an indication of incomplete syntactic control, just as much perhaps as is oversimplified (i.e. short) sentence structure.' Thus, the remaining measures or performance variables to be calculated are:

- 8- The total number of error-free information blocks which measures the learner's success at both syntactic elaboration and error elimination (or avoidance).
- 9- The proportion of error-free information blocks (EFIB) per number of information blocks in the sample and
- 10- finally, the Communicational Capability Index (CCI).

This index is simply a kind of credit system aimed at taking into account the degree of accuracy or correctness manifested by the learner's writing ability. We will return to this point in our discussion, see section 7.10 below.

#### **7.4. Subjects**

The subjects were 103 Zairean students from three different educational institutions in Kinshasa City. The subjects were divided into three groups or proficiency levels according to their institution of origin. Level 1 comprised 57 high school pupils in their final year (i.e. 6th form) or penultimate year (5th form) of secondary education at the Elikya Institute in Kinshasa. By this stage, the pupils have been learning English as a subject-matter for four or three years, respectively. Level 2 consisted of 11 undergraduates from the Faculty of Science (Department of Maths and Physics) at the University of Kinshasa. At the university level, English is taught for Academic Purposes at the rate of two hours per week during the first semester only (i.e. from October to February). Much of the training given at this level focuses on reading skills, and to a lesser extent, on writing short essays. Finally, the third group or level 3 was composed of 35 students in their first and second year as teacher-trainees at the Institut Pédagogique National (IPN) in Kinshasa. These students use English as the medium of instruction and, at the same time, learn English through a variety of courses, including English Grammar, Conversation Classes, English Teaching Practice, etc.

It is obvious that the subjects' level of English training was quantitatively different, in view of the difference in the amount of exposure they were able to obtain within their institutions. The subjects were from varying social and economic backgrounds but none was exposed to English outside the classroom environment. None of the subjects was known to have obvious handicaps either physical or mental (according to their respective tutors). No account was taken of their particular ethnic group. All subjects are Zaireans and French speakers.

### **7.5. Materials and Procedures**

The materials used to obtain writing data consisted of a set of four pictures describing a traffic accident involving a cyclist and a lorry. The pictures used (with not text) were taken from Byrne (1967: 52) Progressive Picture Composition. The subjects were asked to describe what had just happened in the four pictures (A - D). The experimenter told the subjects to describe the story as if they were themselves involved in the accident, that is, from the cyclist's or driver's point of view, rather than limit themselves to talking about the accident as an observer would do. The reason for asking the subjects to describe the situation from the cyclist's or lorry driver's angle was to make them feel involved emotionally and thus, enable them to 'talk' with these visual prompts rather than talk about them. In addition, the subjects were required to write between ten and fifteen lines per picture or episode of the story, and not to worry about mistakes.

The task was conducted in all three institutions by the experimenter himself, during the morning sessions, either the first or second period of the normal classroom timetable, lasting 50-55 minutes each. The time factor was deliberately controlled to avoid excessive self-monitoring on the behalf of the

subjects. The use of dictionaries was not allowed, but instead, the subjects were told that they could ask the experimenter for help regarding the meaning of a single word (not a phrase), if they felt unable to produce an equivalent themselves. Overall, the subjects seemed willing to co-operate with the task. No incident was recorded during the accomplishment of this experiment. Four pupils from the fifth form at the Elikya Institute arrived 20 minutes late; and, although they were asked to get on with the task to avoid disruption of discipline in the classroom, their data were not included in the analysis.

## **7.6. Performance Analysis: An Information Processing Perspective**

### **7.6.1. Method of Analysis of Data**

Our analysis of learner English being reported here was based on writing data obtained from 103 Zairean students at high school, undergraduate and teacher training levels. The students were asked to relate a horrible traffic accident using a set of four pictures as a prompt (Byrne, 1967: 52). The length of the compositions varied considerably in relation to the number of words, both between and within levels of training in English. In order to eliminate the incidence of length of composition in the amount of errors committed, we decided to analyse only the first 100 words of each composition (i.e. a total of 10,300 words for the entire corpus). By dealing with a sample of 100 words per subject it was possible to control for length, topic as well as account for the number of correct and incorrect uses, and possible avoidance of use of a particular structure or feature. The text was then marked with certain symbols in order to be able to identify particular features through computer processing. The actual package used to identify the different errors was the Oxford Concordance Program (Hockey and Marriott, 1980).

### 7.6.2. Error Identification

All errors or incorrect uses related to morphology, syntax and lexis were marked with a ' ' symbol. Those connected with cohesion, contextual relevance and appropriateness, were (impressionistically) subsumed under the cover term 'discourse', and were marked with a '!' symbol. No consideration was given to the traditionally established categories of punctuation, mechanics, etc. After computer processing of data, only those errors which occurred four times or more throughout the entire corpus were regarded as errors. It was assumed that, in theory, if a particular error is being committed by chance, then the probability of its re-occurrence would be 0.33. Thus, an error which appears three or more times in the corpus would have the probability rate of 1.00. In addition, the text was marked with other symbols to enable the identification of sentence boundaries, information blocks, and different clause types. Table 7.1. gives the figures for the number of subjects, words, errors and length of text sample used for the analysis of the compositions.

Table 7.1

Text Sample used for Error Analysis

LEVEL	1	2	3	All Levels
No. of subjects	57	11	35	103
Total no. of words	5,700.	1,100	3,500	10,300
Total no. of errors	327	48	149	524
Length of individual sample in words	100	100	100	100



### 7.6.3. Classification of Errors: Top Ten List

Once identified, the errors were classified into different syntactic and lexical categories. Table 7.2 gives the 'Top Ten' list of the most frequent errors. By and large, the most frequent errors were committed in the area of prepositions (27.1% of all errors), aspect (18.7%), determiners – mainly articles (14.5%), pronouns (both personal and relative (14.1%)) and irregular past tense (11.3%). These five top categories alone accounted for 85.7% of all errors and for just 37.3% of the total output (in words) of the whole corpus. The remaining five categories at the bottom of the list accounted for 14.3% of all errors, and only 4.3% of the total output, i.e. all words used in a 100 word sample.

It also appears from Table 7.2 that as far as these learners are concerned, prepositions and determiners are of great importance, since for every ten words, roughly 1.2 words appear to involve either a preposition or a determiner of some kind (especially articles). In addition, Table 7.3 gives a more detailed account of the errors distribution, i.e. by each of the main sentence constituents, viz noun phrase (NP), verb phrase (VP), and prepositional phrase (PP), as well as level of training. These results (see Table 7.3) indicate that the verb phrase related errors were more frequent (40.5% of errors) than those related to the noun phrase (32.% of errors) or the prepositional phrase (27.1%). Looking more closely at each main constituent, it can be gathered that, regarding the NP constituent, errors involving articles and pronouns were more frequent i.e. these errors accounted for 26% of the total number of errors. Within the VP constituent, errors involving tense and aspect were more frequent (i.e. these accounted for 30% of all errors). 27.1% of all errors were found in the area of NP constituent.

Table 7.2

## Frequent Error Types: Top Ten List

	Total £ of errors	% of errors	£ of uses per 100 wrds	% of uses per 100 wrds
1.) Prepositions	142	27.1	1203	11.7
2.) Aspect	98	18.7	345	3.3
3.) Determiners	76	14.5	1210	11.7
4.) Pronouns	74	14.1	808	7.8
5.) Irregular Past	59	11.3	293	2.8
6.) Auxiliaries	30	5.7	102	1.0
7.) Compound Nouns	20	3.8	200	1.9
8.) Infinitival Comp	12	2.3	14	0.1
9.) Temporal Frame	8	1.5	114	1.1
10.) Subject-verb concord	5	1.0	20	0.2

Table 7.3

## Frequency of Errors Per Major Sentence Constituent and Level

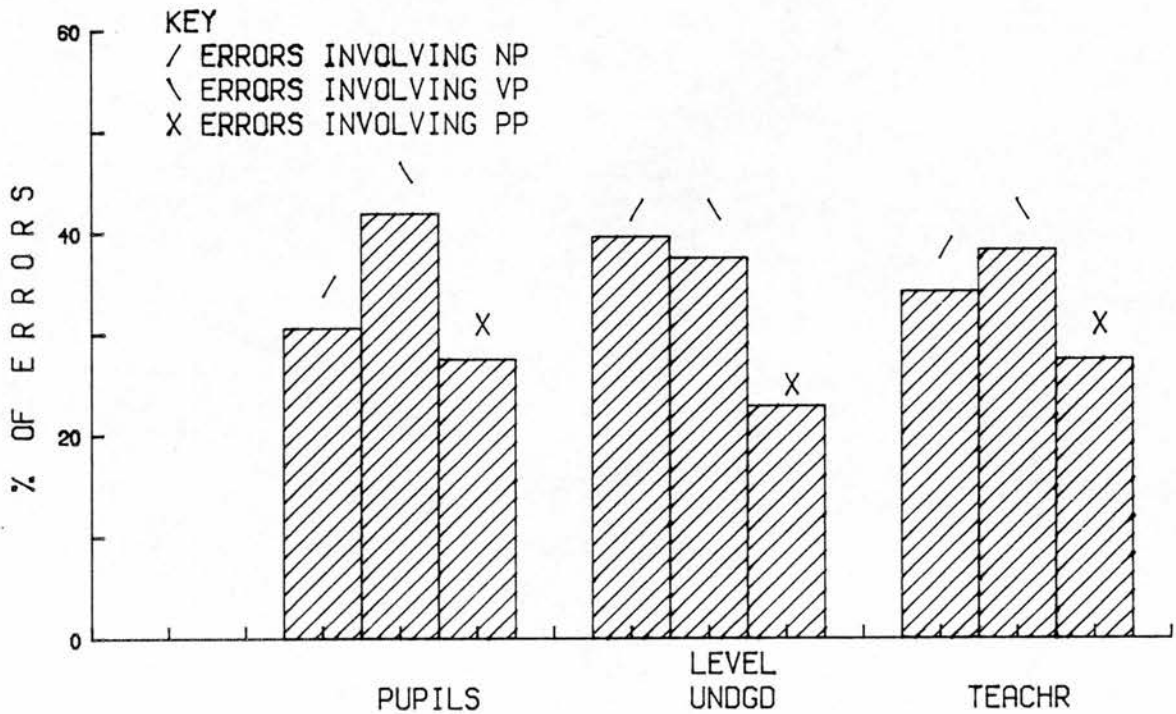
Categories of errors	Level 1		Level 2		Level 3		All Levels	
	n	%	n	%	n	%	n	%
<b>I. Noun Phrase</b> (32.4% of errors)								
1. Definite Article	18	(5.5)	9	(18.8)	19	(12.8)	46	(8.8)
2. Indefinite Art.	10	(3.1)	2	(4.2)	6	(4.0)	18	(3.4)
3. Pers. Pronoun	34	(10.4)	0	(0.0)	11	(7.4)	45	(8.6)
4. Lexical Choice	13	(4.0)	3	(6.3)	4	(2.7)	20	(3.8)
5. Relative Pronoun	20	(6.1)	3	(6.3)	6	(4.0)	29	(5.5)
6. Demonstratives	3	(0.9)	1	(2.1)	3	(2.0)	7	(1.3)
7. Quantifiers	2	(0.6)	1	(2.1)	2	(1.3)	5	(1.0)
<b>II. Verb Phrase</b> (40.5% of errors)								
1. Transitive Verbs	33	(10.1)	10	(20.8)	16	(10.7)	59	(11.3)
2. Intransitive Verbs	70	(21.4)	4	(8.3)	24	(16.1)	98	(18.7)
3. Equational BE	0	(0.0)	1	(2.1)	4	(2.7)	5	(1.0)
4. Auxiliaries	21	(6.4)	1	(2.1)	8	(5.4)	30	(5.7)
5. Temporal Frame	2	(0.6)	1	(2.1)	5	(3.4)	8	(1.5)
6. Infinitival Comp.	11	(3.4)	1	(2.1)	0	(0.0)	12	(2.3)
<b>III. Prepositional Phrase</b> (27.1% of errors)								
Prepositions <i>to, in, of, at, on, by</i>	90	(27.5)	11	(22.9)	41	(27.5)	142	(27.1)
<b>Total # of errors</b>	<b>327</b>		<b>48</b>		<b>149</b>		<b>524</b>	
<b>% of Errors per 100 words</b>	<b>5.7%</b>		<b>4.4%</b>		<b>4.3%</b>		<b>5.1%</b>	
$\chi^2 = 10.58; df = 2; **p < .01$								

Overall, there appears to be a definite tendency for the frequency of errors to decrease as the level of training in English increases. In Figure 7.3, a histogram of the percentage of errors per level and major phrase constituents, reflects this trend. However, this histogram cannot tell us whether this tendency is statistically significant or not. Thus, a Chi-Square was calculated

in order to examine whether the students made fewer and fewer errors as their level of training increased. This trend was found to be statistically significant ( $X^2=10.58$ ,  $df=2$ ;  $p < .01$ ). In other words, the students tended to achieve greater grammatical accuracy or correctness in keeping with their level of training. In addition, this finding is in agreement with that of Larsen-Freeman (1983: 129) who found that on a picture composition task, the number of errors did decline as the level of proficiency increased.

Figure 7.3

Percentage of errors per level for each major  
major sentence constituent



#### 7.6.4. Quantitative Analysis of Data

Let us now turn to the quantitative analysis we conducted in our attempt to find an index of global second language proficiency. These analyses were based on the theory of speech perception and information processing expounded in the sections 7.2. and 7.3 above. First of all, we shall calculate the following variables: (1) the total number of sentences; (2) the total number of information blocks; (3) the total number of clauses; (4) the average number of clauses per sentence; (5) the average number of clauses per information block; (6) the total clause embedding weight; (7) the average clause embedding weight (ACEW); (8) the total number of error-free information blocks; (9) the proportion of error-free information blocks (EFIB). The performance variables which were actually used in the following analyses are (4), (5), (7) and (9) above, since they involved averaged values rather than absolute ones. In order to establish the criterion related validity of these four measures, we conducted a Pearson Product Moment Correlation, including the average number of clauses per sentence (i.e. AVCLAUS), the average number of clauses per information block (AVCLIB), the average clause embedding weight (ACEW), and the proportion of error-free information blocks (EFIB).

Cook (1979: 178) used some of these variables to evaluate the difficulty of 36 writing samples of native speakers of English in terms of information processing. In particular, Cook found a high correlation between the average clause depth (equivalent to our ACEW) and average block length (equivalent to our AVCLIB). The scatter diagram which resulted from his plots showed a smooth curve representing increasing complexity of style. From this finding we feel confident in comparing our Pearson correlation results with those established by Cook (1979). In other words, the latter could be referred to as a criterion.

Table 7.4 gives the results of the correlations between all variables involved. Most of the correlations showed significant results with notable or substantial ( $r$ ) values. First of all, the correlation between the average clause embedding weight (ACEW) with the average number of clauses per information block (AVCLIB) was highly significant ( $r=.87$ ,  $p <.001$ ), thus establishing a similarity between Cook's (1979) results and ours (This accounts for 76% of the total variance). Another set of variables which were of interest to us was the relationship between the ACEW and EFIB (i.e. the proportion of error-free blocks), since we were dealing with second language learners. It was assumed that their communicational capability might be affected by the fact that their linguistic resources are limited. The results show a positive but low correlation between these two variables ( $r = .27$ ,  $p < .002$ ). The lower  $r$  value means that some of those learners who were producing short sentences or simple texts did also commit few errors possibly as a result of the avoidance strategy.

Table 7.4

**Pearson Correlation between 4 Performance Variables from  
Picture Composition Task**

	AVCLAUS	AVCLIB	ACEW	EFIB
AVCLAUS	1.000 (N=0)	0.623** (N=103)	0.540** (N=103)	0.125 (N=103)
	-----			
AVCLIB		1.000 (N=0)	0.871** (N=103)	0.172* (N=103)
		-----		
ACEW			1.000 (N=0)	0.276** (N=103)
			-----	
EFIB				1.000 (N=0)
** p < .01 *p < .05				

The next stage of our quantitative analysis involved an ANOVA to assess the effect of the independent variable (i.e. level of training: high school, undergraduate and teacher training levels) on the dependent variables (i.e. AVCLAUS, AVCLIB, ACEW and EFIB). The analyses were conducted using the BMDP P2V computer programme. First of all, no significance was found between levels and the average number of clauses per sentence. In other words, sheer sentence length could not be referred to for differentiating between the subjects. On the other hand, the ANOVA results (Table 7.5) involving levels and AVCLIB showed highly significant F Values ( $F = 10.18$ ,  $df = 2$ ,  $p < .0001$ ), suggesting that the learners were producing longer information blocks as their level of training increased.

Table 7.5

**ANOVA Results for Average  $\bar{E}$  clauses per Information Block by levels**

LEVEL	N	MEANS	SD
1.) Pupils	57	1.60	0.23
2.) Undergrads	11	1.70	0.44
3.) T-Teachers	35	1.91	0.38
F = 10.18			
** p < .0001			

Appropriate Scheffe tests were conducted to determine which differences between pairs of means were responsible for this F value. From Table 7.6, it appears that learners at level 3 were making significantly much longer information blocks than both level 1 ( $p < .01$ ) and level 2 ( $p < .01$ ). In turn, level 2 were making longer blocks than those at level 1 ( $p < .01$ ).

Table 7.6

**Scheffe Tests - Average  $\bar{E}$  of Clauses per Information Block with Levels**

	LEVEL 2 Undergrads X = 18.7	LEVEL 3 T-teachers X = 66.5	LEVEL 1 Pupils X = 91.2
X = 18.7 Undergrads	-----	** 47.8	** 72.5
X = 66.5 T-Teachers	-----	-----	** 24.7
** p < .01	T crit = 20.32		

The next step was to compare the student's performance in terms of the average clause embedding weight (ACEW), which is a combination of both a



quantitative and qualitative assessment of writing data. The ANOVA results (Table 7.7) were again statistically significant ( $F = 6.88$ ,  $df = 2$ ,  $p < .001$ ), indicating that as the level of training increased, the students were able to achieve greater complexity in their writing ability.

**Table 7.7**

**ANOVA Results for Average Clause Embedding Weight by Levels**

LEVEL	N	MEANS	SD
1.) Pupils	57	1.47	0.18
2.) Undergrads	11	1.58	0.51
3.) T-teachers	35	1.68	0.25
$F = 6.88$ ** $p < .002$			

We regard the ACEW as a truly objective measure for assessing the learner's capability to process information and encode messages in the target language since it is based only on the amount and extent of subordination present in the sample, not on the number of words or erroneous forms within the clauses concerned. In order to specify which differences between pairs of means (or levels) were responsible for the ANOVA F-value, we conducted Scheffe tests whose results are shown in Table 7.8.

Table 7.8

## Scheffe Test - Average Clause Embedding Weight with Levels

	LEVEL 2	LEVEL 3	LEVEL 1
	X = 17.35	X = 58.63	X = 84.02
X = 17.35	-----	** 41.28	** 66.67
X = 58.63	-----	-----	** 25.39
**p < .01 T Crit = 16.51			

According to these results, those students who were at a more advanced level were also significantly, more capable of producing more complex or elaborate text in the target language. In addition, the ACEW measure is, in our view, a highly valid measure to assess second language proficiency in general, and the degree of risk-taking that the students were willing to initiate. Since the respondents were told not to worry about making mistakes during the accomplishment of this task, it can be hypothesized that they took as many opportunities as possible to focus on the content or meaning rather than grammatical accuracy. Thus, ACEW would reflect the maximum of risk-taking as well as the minimum of avoidance on the respondent's behalf, since it does not take errors into account. In other words, those respondents who would be more inclined to avoid running into trouble would also tend to write short, simple sentences, thus would have a low score on this measure (i.e. ACEW). Since the above results indicate that the ACEW values increased in keeping with the level of training, we can feel confident in saying that, overall, the more advanced learners were also the ones who were taking more risks. From a purely pedagogic viewpoint, this tendency could well be encouraged.

Let us now turn to the analysis of variance involving levels of training and the proportion of error-free information in the sample. These ANOVA results are given in Table 7.9. The results were again highly significant. ( $F = 8.90$ ,  $df = 2$ ,  $p < .0003$ ). This meant that the learners were again improving their ability to write compositions which contained fewer and fewer grammatical errors, as they progressed in English.

Table 7.9

## ANOVA Results for Error-free Information Blocks by Level

LEVEL	N	MEANS	SD
1.) Pupils	57	0.31	0.16
2.) Undergrads	11	0.49	0.20
3.) T-Teachers	35	0.38	0.23
F = 8.90 ** p < .0003			

During our classification of errors, we established (by  $\chi^2$ ) that the more advanced the learners were, the less frequently they made grammatical errors. But this was an incomplete picture of the learners performance since it does not take into account the features of their output which were error-free. In addition, at that stage, we were dealing only with frequencies, rather than the amount of the differences that could be found in the data. In order to detect such differences between pairs of levels, we conducted post hoc comparisons by the Scheffe method. These results were not significant. This means that these 3 groups of learners were different in terms of grammatical correctness but their differences were not big enough to reach statistical significance. We speculate that those who were at a more advanced stage did write both

elaborate (i.e. risk-taking strategy) and grammatically accurate information blocks; but at the same time, those who were at a less advanced stage were writing short information blocks (i.e. risk avoidance strategy) and, proportionately, reduced their chances of making errors.

In order to obtain a clearer picture regarding the possible role played by the strategy of avoidance, we decided to examine whether the groups could be differentiated in terms of grammatical correctness, provided that the degree of risk-taking (or ACEW) is controlled for. In other words, we asked ourselves the following question: 'Did error-free behaviour change per level when language competence (as reflected by risk-taking) was kept constant?' In statistical terms, this would mean that the proportion of error-free information blocks is the dependent variable, level the independent, and the average clause embedding weight is the covariant. Thus, the null hypothesis could be stated as follows: Assuming that every student was taking the same amount of risks, there was no significant difference between the 3 groups as regards error-free performance.

We then conducted an ANOVA with covariance (using BMDP P2V Subprogramme) to see if correctness was due to level of training in the hypothetical situation where the students' degree of risk-taking was equivalent throughout the sample. The ANOVA results are given in Table 7.10.

**Table 7.10**

**ANOVA Results for Error Free Information Blocks by levels with ave. Clause Embedding Weight, ACEW, as Covariant**

SOURCE	SS	df	MS	F	P
Levels	0.433	2	0.216	6.08	0.003***
Covariant: ACEW	0.112	1	0.112	3.15	0.079
** p < .01					

These results are significant ( $F= 6.08$ ,  $df = 2$ ;  $p < .003$ ) which suggests that correctness did change per level when competence (or risk-taking) was assumed to be equal. However, appropriate Scheffe tests to determine which pairs of levels were responsible for these differences failed to reach statistical significance. This indicates that, although the students were encouraged to take risks (i.e. not to bother about making errors as the experimenter told them) they still adhered to the well established criterion in most classroom situations, viz, to avoid making errors. One way of achieving this would be to write simple sentences. In other words, those at an advanced level are producing few errors while writing more complex sentences. However, those at a lower level, or at least many among them seem determined to produce few errors; and they do so by writing simple, short information blocks.

### **7.7. Learner Language Behaviour: A Re-appraisal**

Is there an error in Error Analysis?

In the previous section, we were more concerned with the analysis of the learner's performance with a view to establish a global index of development. Our analytical technique was the clause or cluster of clauses as a unit of information. In the present section we are going to focus on the learner's performance as an Interlanguage. Strictly speaking, such an analysis ought to be conducted without any attempt to relate his behaviour to other types of performance (e.g. to the learner's  $L_1$  performance, or to the target language, or, indeed to other  $L_2$  learners). In other words, the learner's performance ought to be regarded as a separate norm in its own right (Selinker, 1972); and, thus,

there are no 'errors' in such a performance.

However, in real terms, the notion of error is somewhat indispensable, especially in an educational situation where norms are usually of some concern. In the previous section, for instance, we observed that this norm-awareness can affect the learner's performance (that is, he may set himself grammatical accuracy as a norm and thus adopt the risk-avoidance strategy to reduce the chances of making errors). Similarly, early error analyses were mainly a reflection of the norm-relatedness of learner's performance. Consequently, they focused on the 'erroneous' aspect of learner language and often ignored its 'non-erroneous' part. In the mid 70's, however, some researchers began to undertake analyses of both errors and non-errors (Zydatiss 1974, 1976). Nevertheless, a more complete picture of the learners' performance could not be obtained until another stage in error analysis could be reached, viz, the investigation of causes of errors. In this connection, too, it is easy to understand why learner strategies were, historically, investigated later. The study of the causes of errors involved the researchers moving away from simple error identification, to classification into categories, and, then, to the explanation of error causes, whence to the examination of learner's strategies (Stenson, 1975; Kasper, 1982; Beebe, 1980, 1983).

Of course error causes can be rather elusive to pin down. In principle, this would involve an account of both errors and non-errors (i.e. all the attempts to use a certain form or feature, as well as possible 'non-attempts' to use the form or feature (i.e. all possible instances of risk-taking or avoidance). To my knowledge, such studies are still to be undertaken.

Usually, there are two major dimensions along which errors are explained. The first distinction is between learner-internal and learner-external causes

(Faerch, Haastrup and Phillipson, 1984). Errors may be the result of internal cognitive procedures, e.g. transfer from  $L_1$  or overgeneralization to new contexts of Interlanguage rules. Alternatively, errors may be the result of factors external to the learner such as inadequate teaching materials or practice. The second way of explaining errors is to distinguish between direct and indirect causes. For instance, the incorrect use of the definite article *the* frequently made by Zairean learners of English may be caused by overgeneralization, a learner-internal factor, as in e.g. *\*It is the Eve's birthday* (i.e. It was Eve's birthday.) Here the direct cause is a cognitive one whereby the use of the article *the* is generalized to all NP's including those already marked for definite reference by means of the possessive case. The indirect cause might be due to teaching presentation and practice whereby particular forms or structures are taught in isolation or independently from their functions. In addition, other factors can affect indirectly the learner's performance and account for 'errors' or indeed 'non-errors'.

One such factor is the learner's psychological traits, e.g. the willingness to take linguistic risks rather than paying attention to form or accuracy. Another factor could be the learner's assumptions about which parts of their  $L_1$  can be freely transferred to their IL (Kellerman, 1977). With some learners, it may be a rather 'conscious' or sophisticated strategy toward  $L_2$  approximation. This is what Zydattiss (1976: 361) has referred to as 'over-compensation', the rationale being something like 'if it is not  $L_1$ , it is probably correct'.

However, there is also the problem of causal ambiguity in that errors can often have more than one direct cause, i.e. transfer and overgeneralization can function together. For instance, when some of the Zairean learners (or French-speakers learning English) produce compound noun phrases, such as *He*

went to center town instead of *He was going to the city centre*, this may be due to transfer from the French (i.e. the equivalent in French is *Il allait vers le centre ville*). Equally, it may be due to the generalization of the English rule for compound noun formation. In other words, the learner may have become linguistically aware that in order to form compound nouns in English, one has to apply a rule which has two stages: 1) to delete both the preposition and second instance of *the*, from the phrase *the centre of the city*; 2) to invert the word order of both nouns. Thus, by producing the incorrect form *\*centre town*, the learner may have applied only one stage of the rule. Owing to this insufficient linguistic knowledge, the learner, then, resorts to the French word-order to make up for stage two of this rule. Thus, he produces the form *\*centre town* or *\*centre city*, like in French and instead of *city centre*.

Until recently, the standard error analysis view was still based on the assumption that, if one had a more direct access to the psycholinguistic processes, one could allocate errors to one of the two internal causes. But as shown by the examples given above, both transfer and overgeneralization can operate at the same time. A consequence of this observation is that, if we attempted to quantify exactly the proportion of errors due to one cause rather than the other, we would, in effect, trivialize the whole effort to investigate the complex role played by the cognitive processes in SLA, which is precisely what some of the earlier IL studies set out to do (Lamendella 1977, 1979; Selinker and Lamendella 1978). So, to answer the question we asked ourselves at the beginning of the present section: Yes, there may still be an error in error analysis. And what do we do about this?

A more appropriate way of looking at the learner's performance as an IL would be to try and characterize, qualitatively, the intricate ways in which the



various errors and causes or errors interact with each other. In addition, from the point of view of L<sub>2</sub> pedagogy, error analysis as outlined here would serve as a tool for discovering the external causes of errors, as these may provide invaluable information leading to a more felicitous teaching - learning interaction, e.g. syllabus negotiation, richer comprehensible input, a more individualized type of language learning, etc. In the next section, we will present a detailed account of learner language behaviour.

In sections 7.6.2 and 7.6.3 the errors were classified into different grammatical categories. We drew the top ten list of the most frequent errors (Table 7.2) and we then established the errors distribution according to the three main sentence constituents, viz, the Noun Phrase (NP), Verb Phrase (VP) and the Prepositional Phrase constituents (PP), (Table 7.3). In the following section, we will scrutinise closely these different errors and their possible causes. We will look at those errors committed in the NP area first, then those in the VP, and finally, those related to the PP, following the same order of presentation as in Table 7.3. Each example will be followed by two numbers in parentheses referring to the respondent's level and identification number, respectively, to enable us to relocate the origin of the examples in the corpus.

## **7.8. Learner Language Behaviour: Noun Phrase Related Features**

Both errors and non-errors for the NP constituent are given in Table 7.11.

### **7.8.1. Definite Article**

Let us first look at some of typical error examples.

- (1) \*I am going to deal with the description of the circulation accident.  
(I am going to describe a traffic accident)
- (2) \*I took my bicycle to go the service (3, 39).  
(I took my bicycle to go to work)

- (3) \*I began to lose the control (3, 18).  
(I began to lose control)
- (4) \*It is the Eve's birthday (2, 96).  
(It was on Eve's birthday)
- (5) \*He was upset by an car which bring the cases of the ailments (1, 106).  
(He was knocked down by a lorry which was carrying (some) crates containing food stuffs).
- (6) \*In the morning, I went to the school (1, 118).  
(In the morning, I went to school).

Most errors involving definite articles are due to the oversuppliance of *the* where *0* ought to be used in Standard English, i.e. generalization of a TL rule. But it is also possible to attribute these errors to a secondary cause, i.e. to the interference of French. In other words, the learners may be having difficulty in delineating the appropriate uses or function of *the*, and so they rely on the syntax of French (i.e. Transfer) where the mass-count nouns distinction is grammatically neutralized (e.g. in (3) and (4)). Quite apart from the overgeneralization of the use of *the* to mass NPs contexts, there still is, among many learners, some confusion as to when to use definite rather than indefinite reference (e.g. (1) and (5)). In other words, the function of the definite article *the* seems to be extended by some learners to identify not only specific referents that are known to the hearer, but also those that are not known to the hearer, that is, *the* is extended to the function of an indefinite article (e.g. in (1) . . . *the circulation accident*). Sometimes the function of *the* is overstretched (i.e. when *the* is used) to identify NPs that are already definite, e.g. by way of the possessive case as in (4). Table 7.11 gives the absolute numbers (n) and percentages of NP related 'errors' (\*), 'non-errors' (+) as well as the total number and percentages of attempts (A) to use particular forms within the 100 word sample. Looking at the 3 right-most columns indicating total number and percentages, it can be seen that 8.8% of all errors committed in the whole

Table 7.11

## Noun Phrase Performance Features

Features	Level 1			Level 2			Level 3			Total		
	+	*	A	+	*	A	+	*	A	+	*	A
<b>I NOUN PHRASE</b>												
<b>1.) Definite Article</b>												
<i>the</i>												
n=	397	18	415	94	9	103	250	19	269	741	46	787
%=	19.3	5.5	17.4	21.7	18.8	21.4	19.3	12.8	18.7	19.6	8.8	18.3
<b>2.) Indefinite Article</b>												
<i>an</i>												
n=	159	10	169	40	2	42	105	6	111	304	18	322
%=	7.7	3.1	7.1	9.2	4.2	8.7	8.1	4.0	7.7	8.0	3.4	7.5
<b>3.) Personal pronoun and Possessive Adjectives</b>												
<i>he</i>												
n=	302	8	310	57	0	57	98	2	100	457	10	467
%=	14.7	2.4	13.0	13.2	0	11.9	7.6	1.3	6.9	12.1	1.9	10.8
<i>him</i>												
n=	47	3	50	11	0	11	22	4	26	80	7	87
%=	2.3	0.9	2.1	2.5	0	2.3	1.7	2.7	1.8	2.1	1.3	2.0
<i>it</i>												
n=	38	3	41	7	0	7	33	5	38	78	8	86
%=	1.8	0.9	1.7	1.6	0	1.5	2.6	3.4	2.6	2.1	1.5	2.0
<i>her</i>												
n=	4	20	24	2	0	2	1	0	1	7	20	27
%=	0.2	6.1	1.0	0.5	0	0.4	0.1	0	0.1	0.2	3.8	0.6
<b>4.) Lexical Choice</b>												
<i>center</i>												
n=	4	1	5	0	3	3	1	0	1	5	4	9
%=	0.2	0.3	0.2	0	6.3	0.6	0.1	0	0.1	0.1	0.8	0.2
<i>lorrie</i>												
n=	25	7	32	5	0	5	42	0	42	72	7	79
%=	1.2	2.1	1.3	1.2	0	1.0	3.3	0	2.9	1.9	1.3	1.8
<i>man</i>												
n=	56	5	61	9	0	9	38	4	42	103	9	112
%=	2.7	1.5	2.6	2.1	0	1.9	2.9	2.7	2.9	2.7	1.7	2.6
<b>5.) Relative Pronouns</b>												
<i>who/which</i>												
n=	61	20	81	16	3	19	35	6	41	112	29	141
%=	3.0	6.1	3.4	3.7	6.3	4.0	2.7	4.0	2.8	3.0	5.5	3.3
<b>6.) Demonstratives</b>												
n=	33	3	36	7	1	8	25	3	28	65	7	72
%=	1.6	0.9	1.5	1.6	2.1	1.7	1.9	2.0	1.9	1.7	1.3	1.7
<b>7.) Quantifiers</b>												
<i>Much/many</i>												
n=	16	2	18	0	1	1	8	2	10	24	5	29
%=	0.8	0.6	0.8	0	2.1	0.2	0.6	1.3	0.7	0.6	1.0	0.7
<b>Total for Noun Phrase</b>												
n=	1142	100	1242	248	19	267	658	51	709	2048	170	2218
%=	55.4	30.6	52.0	57.3	39.6	55.5	50.9	34.2	49.2	54.1	32.4	51.5

Key:

n = number of occurrences  
+ = correct use  
\* = incorrect use  
A = frequency of attempts to use form  
% = row percentage

sample were found to involve the definite article. This trend of overuse of *the* is apparent at all three levels, but it is more pronounced at level 2 (18.8%), where *the Eve's birthday* type of error is common but hardly exists anywhere else.

In the meantime, the percentage of all the correct uses of *the* is roughly the same at all 3 levels (i.e. 19.3% at levels 1 and 3 and 21.7% at level 2). In addition, level 1 appears to be committing fewer errors involving the article *the*, 5.5% at level 1, but up to 12.8% at level 3 and even 18.8% at level 2. Since the most frequent errors (as seen in the examples above) derive from the overuse of *the*, we can then conclude that overgeneralization is the main strategy being used by the more advanced respondents (i.e. levels 2 and 3). On the other hand, those learners who are at level 1 may be using a combination of both transfer and generalization. This is what seems to be going on in the examples (5) and (6) above, where there are instances of *an* and *the* which may be due to either the phonology of French or both the syntax of French (i.e. transfer of the function of *the*) and that of English (i.e. overgeneralization of *the* to the wrong contexts) as in:

(7) \*He was upset by an car . . . (1, 106)  
? (Il etait reverse par une auto.)

(8) \*In the morning I went to school (1, 108)  
(Le matin j'allais a l'ecole.)

### 7.8.2. Indefinite Article

Here are some of the most typical examples in the corpus.

(9) \*He went there by a bicycle. (1, 130)  
(He went there by bicycle)

(10) \*The history is about a accident. (1, 127)  
(The story is about an accident.)

- (11) \*. . .he didn't see a car who crossed this road. (2, 80)  
 (he couldn't see the lorry which was crossing the road.)
- (12) \*I never seen in my life a kind of driver. (3, 26)  
 (In my experience, I have never come across a driver  
 of this kind)
- (13) \*It was an horrible accident. (1, 156)  
 (It was a horrible accident.)

Many errors here are still committed at the phonological level owing to a certain confusion between when to use *a* and *an* (see (10) and (13) above). This type of error is more frequent at level 1 where the correct use of this form seems to be in the pre-systematic stage. Another frequent error found at this level is the *go by a bicycle* type in example (9). Here, the indefinite article is overgeneralized to phrases such as this where *0* is required, viz. generic reference i.e. non-definite, non-specific NPs, but known to hearer, rather being used in this context of non-definite, specific NPs not known to the hearer. Elsewhere, as in (11) above, the indefinite article *a* is wrongly used instead of *the* to mark an NP which is postmodified by a relative clause. The overall trend across the 3 levels of training indicates that the more advanced students are committing errors involving *a* more frequently than level 1 (see Table 7.11). Again, this may be due to a greater willingness or ability to produce more elaborate sentences and thus run into trouble more frequently. This willingness to produce complex sentences can be seen in example (12), which is a rather bold attempt to handle quite substantial chunks of information, as indicated in the more target-like sentence given in brackets under (12). In this example, produced by a student at level 3, *a* has been generalized to the function of the demonstrative *this*.

### 7.8.3. Personal Pronouns

Errors involving personal pronouns are by far the most frequent with the NP constituent. In Table 7.11, these errors account for 10.4% of all errors committed at level 1, and 7.4% at level 3. Surprisingly enough at level 2, no error involving personal pronouns was found (see Tables 7.3 and 7.11). Table 7.11 shows that this all correct use of personal pronouns is not due to the strategy of avoidance, since there are at least as many attempts made by level 2 respondents as those by other levels. Presumably, this may point to a distinctive feature of strategic competence at this level. Let us look at the most common errors involved in these structures. The pronouns encountered in the corpus were *he*, *him*, *it* and *her* in decreasing number of occurrences.

- (14) \*When he arrive late, the headmaster punishes her. (1, 112)  
(Whenever he arrived late, the headmaster punished him.)
- (15) \*The police stop he and they came, they sought the ambulance. (1, 101)  
(The police stopped him and came over, they called an ambulance?)
- (16) \*As he was not looking in front him. . . (3, 67)  
(As he was not looking in front of him. . .)
- (17) \*Before the bicycle it has a lorry which comes. (3, 73)  
(Ahead of the bicycle, there is a lorry. . .?)
- (18) \*This street it was at my left. (3, 15)  
(This street was on my left.)
- (19) \*There was in Paris one young man who liked his bicycle very much. One day his father told him to buy a car for him but the young man refused. (1, 142)  
( . . . his father told him that he wanted to but him a car).

The main learning problem in this area seems to involve gender and case marking (e.g. in (14) and (15)). There is also the tendency to use the expletive forms *it* and *there is* interchangeably, presumably as a result of the transliteration of the French equivalent *il y a* ( i.e. *there is* or *it is*, as in (17)). In addition, in example (19), we can see that the learner is confronted with the

problem of co-reference, that is, there is some confusion regarding the identification of the correct agent NP or subject of the complement clause - *to buy a car for him*.

One of the most frequent errors committed by respondents at level 1 was the wrong gender assignment using the feminine possessive adjective form *her* instead of *his* mostly in the context of NPs, which in French are marked for feminine regardless of whether the possessor is male or female in English. For example,

- (20) \*After her lunch, he went to her bedroom and he take her bicycle. (1, 112)  
(After having breakfast, he went to his bedroom and took his bicycle.)
- (21) \*Charles father is a very hard worker in a big company. Her mother is teaching chemistry in I.S.P., his grandfather is leaving in Lubumbashi. (1, 111)  
(Charles' father is a dedicated worker in a big company. His mother teaches chemistry at the ISP Teacher's college, his grandfather lives in Lumbumbashi.)
- (22) \*One man goes to the country to look for her family. . . he drives her bicycle very quickly. . . he isn't see the car before him, he thinks to her mother her sister who hadn't seen five years ago. (1, 132)  
(He is going to the country to visit his relatives. . . he is cycling very fast. . . he can't see the car in front of him, he is thinking of his mother, his sister whom he hasn't seen for five years.)

All of these errors involving gender assignment were found at level 1 only (6.1% of all errors at this level). Although the form *her* occurred quite infrequently at both level 2 and 3, we cannot say that they were being avoided altogether since they are actually used at these two levels, correctly, all the time. In connection with level 1, however, we may even be tempted to think that the learners do not even know that the masculine *his* exists. In example (20), *her* is used with both masculine and feminine NPs in French, viz. *her lunch*

(i.e. *son déjeuner*) or *the bedroom* (i.e. *sa salle à coucher*). In this case, the strategy which is being used is: for the 3rd person singular, the pronoun is *he* and the possessive adjective is *he + r* (i.e. overgeneralization). However, when we look at the examples (21) and (22), we realize that both forms (i.e. *his* and *her*) have been learned; but they are still at a pre-systematic stage whereby they are used by the same learner(s) as free-variants of the same form. In both of these examples ((21) and (22)) there is clear evidence that, indeed, transfer from French is the main cause of the incorrect uses of *her*. In French, the possessive adjective agrees with the gender of the NP that it modifies, not the gender of the possessor as is the case in English. Thus, in example (21), we have both *Her mother is teaching* and *his grandfather is leaving*. In both cases, Charles is the antecedent of *her* and *his*; but these forms are not in agreement with their antecedent's gender, but that of their immediate NPs, which is exactly the correct way of using anaphoric reference in French. This feature of the Interlanguage of these learners necessitates particular attention on behalf of materials designers or those involved with preparing pedagogic grammar for French-speaking learners of English, especially in the beginning stages.

#### 7.8.4. Lexical Choice

Errors subsumed under this broad category include mainly instances of compound nouns and word order. Most errors are in connection with the words *\*centre town*, *\*driver lorry*, *\*Bicycle man* as in the following examples:

- (23) \*. . .he went to buy something in the center town. (2, 77)  
(He was going to the city centre to buy something.)
- (24) \*Sometime letter the driver lorrie climbed up to see  
who fell. (1, 133)  
(Some time later, the lorry driver came down to see whether  
anything serious had happened to the cyclist. . .)
- (25) \*It was a big accident between a lorry car and bicycle. (1, 131)  
(It's about an accident involving a lorry and a bicycle.)



(26) \*There is a bicycle man who arrive. . . (3, 39)  
 (A cyclist is on his way. . .)

(27) \*He became man and began to work. (1, 142)  
 (When he became an adult, he started working.)

Errors involving wrong word-order in compound nouns are quite common among French-speaking learners of English. These learners often show incomplete mastery of the TL rule in this connection. This rule necessitates going through two different steps: first, to drop both the preposition *of* and the definite article from say *centre of the city*; second, to invert the two NPs word-order (i.e. *city centre*). Some learners seem to be aware of this TL rule which differs from French, but many others seem to be unable to apply it, in part or in full, if they are actually aware of it. Thus, to a great extent, the French word-order prevails in many cases, but the preposition and article are dropped.

In addition, sometimes learners produce compound nouns which are redundant in English (as in *a lorry car*). This happens when the learner feels uncertain whether the chosen word is the right one, and so the learner resort to the strategy of 'over-compensation' (Zydattiss, 1976), meaning something like <if it is not  $L_1$  then it is correct>. In a sense, too, the learner is adopting some kind of over-kill strategy whereby he ensures that, by using both the hyperordinate and superordinate terms, he must get the right choice at all costs. Judicious inclusion in the teaching materials of such notions as hyponymy, antonymy, synonymy, etc., would seem helpful for most learners at all three levels. The need for such notional awareness manifests itself further in examples (26) and (27) where vague terms such as *bicycle man* and *man* are used to make up for the lack of the more precise terms, viz *cyclist* and *adult* respectively. If we look at Table 7.11, we realize that errors involving *man* (i.e.

use of a vague cover term, or superordinate) were more frequent (1.7% of all errors), followed by those related to *lorry* (1.3%, especially over-compensation) and, then, only word-order, as in \**centre town* (0.8%, probably due to transfer from the French equivalent *centre-ville*).

### 7.8.5. Relative Pronouns

These types of errors are among the most frequent ones committed in connection with the NP constituent, i.e. with 5.5% of all errors, they occupy the third position after the definite articles and personal pronouns. And yet, only two types of relative pronouns are found in the data, viz. *who* and *which*. No instance of the other forms are found (e.g. no *whom*, *whose*, etc.). Let us look at some of the most typical incorrect uses involving *who* and *which*.

- (28) \*But it wasn't my fault because I saw a lorry who is driven by a mad man. (3, 28)  
(. . . I saw a lorry which was being driven by a mad man.)
- (29) \*The man who is going to the office which don't see he is late in the office. (1, 125)  
(The man who was going to the office, didn't realize that he was getting there late.)
- (30) \*This bicycle is the gift who my father gave me. (3, 16)  
(This bicycle is a gift that my father gave me.)
- (31) \*He didn't see a car who crossed the road. (2, 80)  
(He couldn't see the lorry which was crossing the road.)
- (32) \*This is the story which the Picture A. (1, 112)  
(This is the story which is behind the first picture.)

In the corpus, the percentage of non-errors for both *who* and *which* is roughly the same (i.e. 1.8% for *who* and 1.2% for *which*; total is 3% as shown in Table 7.11). However, the percentage of errors is substantially different (i.e. 4.6% for errors involving *who* and only 0.9% for *which*). This would seem to suggest that *which* is the basic unmarked form for these learners and *who* is the marked form, thus, the more difficult one. Alternatively, this may mean that *which* is

being used as a relative pronoun and only *who* is used to fulfil the function of both *which* and *who*. If the first suggestion were true, (that is, if *which* was the unmarked form) then we would find instances of *which* used by learners in a pre-systematic stage for both human and non-human NPs. However, no such instances are present in the corpus since *which* is used with non-human NPs.

What we do find instead, are many examples such as (28), (30), and (31) above where *who* is used incorrectly, i.e. overgeneralized to refer to non-human NPs while being used correctly (60% of the non-errors) with human NPs. As far as the other alternative is concerned (i.e. whether *which* is not regarded as a relative pronoun but only *who* is so) we are less certain to give a clear-cut yes/no answer. Looking at the errors alone would in fact favour this proposal. For instance, in (29) above, the form *which* is actually used to fulfil the role of the complementizer *that*, whereas *who* is used in the appropriate way. In example (32) we have a slightly ambiguous sentence, but, with the benefit of doubt being given to the learner, we can say that here *is* is used in a normal sense too. In this case, then, *who* is unmarked and better understood by these learners than *which*, but both are still in the process of being mastered.

#### 7.8.6. Demonstratives

A few of these errors are due to the learners' failure to mark for the plural, probably as a result of inadequate teaching presentation (at the phonological level) of the sound distinction between *this* /ðɪs/ and *these* /ði:z/, or between *this* /ðɪs/ and *his* /hɪz/.

- (32) \*In this pictures we have four parts. (3, 57)  
(These pictures are grouped into four parts.)
- (33) \*Once upon a time a man going to the office. This name  
is Diasso (1, 126)  
(One day, a man called Diasso was going to work.)

Other errors involve a certain confusion between *this* and forms such as *there* and *that* but are relatively infrequent in the corpus as a whole (i.e. only 1.3% of all errors). Some of these errors are found in the following examples:

- (34) \*He did not put attention that at the crossroads this was a big car. . . (2, 95)  
(He didn't realize that there was a lorry going around the crossroads.)
- (35) \*Then this day, he had some problems with his wife.  
(Then on that day, he had an argument with his wife.)

### 7.8.7. Quantifiers

Errors involving *much/many* are infrequent (only 1% of the total) in relation to the whole sample. However, they seem to cause greater difficulty at the more advanced levels (2.1% at level 2 and 1.3% at level 3) than at level 1 (only 0.6%, i.e. one third of those found at level 2). Some of the examples include:

- (36) \*The man had much ideas in his mind. (1, 143)  
(The man had many ideas in his mind.)
- (37) \*I do many efforts to go away . . . (3, 4)  
(I made a great effort to leave.)
- (38) \* . . .as the way is longer, I take many time. (1, 129)  
(I had a long way to go, so it took me a long time.)

The results in Table 7.11 suggest that these quantifiers do not constitute a major learning problem. The learners generally know to handle these forms, even though some aspects of English syntax remain to be fully mastered. In general, though, these quantifiers operate in a way similar to that of their French equivalent (i.e. *beaucoup de* for *much* and *beaucoup des* for *many*). Therefore, transfer may have a facilitating effect here; except in the case of some borderline uses between grammar and idiom, such as *to make a great effort*, or *a lot of time* as opposed to *many times*.

## 7.9. Learner Language Behaviour: Verb Phrase Related Features

The analysis of Verb Phrase related features concentrated on five grammatical categories, viz 1) transitive verbs; 2) intransitive verbs; 3) copula *je g*; 4) auxiliaries; and 5) subordination (including temporal frame and complementation). Using the same criteria for error identification established earlier (section 7.6.2) above, we obtained the distinction given in Table 7.12. As was the case regarding the NP, we took into account both 'errors' (\*) and 'non-error' (+), as well as the total number of attempts (A) to enable us to discover possible cases of 'avoidance' of particular forms or features.

### 7.9.1. Transitive verbs

Verbs in this category include *leave, ride, see, take, and have* (i.e in the sense of *have got* or *have in one's possession*, in other words, as a relational verb (Quirk et al. 1972: 96). Most of the errors committed in this area involve the irregular past tense, and to a lesser extent the problem of establishing the temporal frame for durational happenings.<sup>1</sup> Let us look then at some of the most frequent type of incorrect constructions involving each of these verbs.

#### Leave

- (1) \*Last Monday, he leaved their house at 35 past seven. (2, 94)  
(Last Monday he left home at 7:35)
- (2) \*He left in a small house near the big market. (3, 56)  
(He (one of my friends) lived in a small house.)
- (3) \*His grandfather is leaving in Lubumbashi. (1, 111)  
(His grandfather lives in Lubumbashi.)

As can be seen from these incorrect uses, many errors are due to a certain

---

<sup>1</sup>The issue of temporal frame will be dealt with later on.

Table 7.12

## Verb Phrase Performance Features

Features	Level 1			Level 2			Level 3			Total			
	+	*	A	+	*	A	+	*	A	+	*	A	
II. VERB PHRASE													
1.) Transitive verbs													
<i>leave</i>	n=	0	2	2	2	1	3	1	4	5	3	7	10
	%=	0	0.6	0.1	0.5	2.1	0.6	0.1	2.7	0.3	0.1	1.3	0.2
<i>ride</i>	n=	17	10	27	10	5	15	20	4	24	47	19	66
	%=	0.8	3.1	1.1	2.3	10.4	3.1	1.5	2.7	1.7	1.2	3.6	1.5
<i>see</i>	n=	73	12	85	4	0	4	28	1	29	105	13	118
	%=	3.5	3.7	3.6	0.9	0	0.8	2.2	0.7	2.0	2.8	2.5	2.7
<i>take</i>	n=	21	5	26	5	4	9	27	4	31	53	13	66
	%=	1.0	1.5	1.1	1.2	8.3	1.9	2.1	2.7	2.2	1.4	2.5	1.5
<i>have</i>	n=	11	4	15	0	0	0	15	3	18	26	7	33
	%=	0.5	1.2	0.6	0	0	0	1.2	2.0	1.2	0.7	1.3	0.8
<b>Total</b>	n=	122	33	155	21	10	31	91	16	107	234	59	293
	%=	5.9	10.1	6.5	4.8	20.8	6.4	7.0	10.7	7.4	6.2	11.3	6.8

Key:

- n = number of occurrences  
 + = correct use  
 \* = incorrect use  
 A = frequency of attempts to use form  
 % = row percentage

confusion between two different verbs, viz *leave* (whose past tense is *left*), and *live* (whose past tense is *lived*). Clearly, the learner's knowledge of the irregular past tense is still in the pre-systematic stage, i.e. sometimes the form of *leave* is matched with its meaning, as in (1), and sometimes it is not, as in (2) and (3). The problem with the obvious confusion between *leave* and *live* (as in (2) and (3)) may have originated from inadequate treatment of the sound discrimination by the teaching techniques to which they had been exposed. Looking at the results in Table 7.12, we can see that all of the attempts made, at level 1, were unsuccessful. However, two out of the three attempts at level 2, were correct; whereas, at level 3, only one out of five attempts was correct. The performance of level 2 respondents is somewhat misleading, though, because when the form *left* is correctly used at this level, it is mainly as a noun for location (i.e. the opposite of *right*) and in idiomatic phrases such as *there are fifty minutes left to reach the school*, (2, 91). Indeed, when we examine the corpus more closely we realise that all levels attempt to use *left* in this type of construction more frequently than in the usual sense of a verb involving a 'transitional event' (Leech 1971: 19).

### Ride

Most errors involving *ride* are related to the use of the progressive aspect as well as the irregular past. For example,

- (4) \*A man who ride the bicycle can't see this car and goes on.  
(1, 163)  
(The man, who was riding a bicycle, couldn't see the lorry,  
and so he went on riding.)
- (5) \*When he rode, he didn't pay attention in front of him. (2, 77)  
(While riding he didn't look ahead. . .)

In Table 7.12, the relative frequency of errors at all three levels is given. Level

2 has a greater amount of errors (10.4%) than both levels 1 (3.1%) and level 3 (2.7%). When we examine the correct uses in the text, we find that almost three out of five correct forms (i.e. 30 out of 47) involve the past progressive; only two instances of errors are related to this subcategory. On the other hand, irregular past remains the main learning problem, especially at the more advanced levels (2 and 3). Many errors at these levels are due to the overgeneralization of *rode* to constructions where the progressive would be required in the target language as in: \**he rode on his bicycle* (He was riding his bicycle).

### See

Most errors in connection with the verb *see* occurred in constructions where it is used in conjunction with modals, as in

- (6) \*In another direction a fat car came but Mr. Malu don't saw. (1, 156)  
(A big car had just appeared from the other direction but Mr. Malu didn't see it.)
- (7) \*I run as fast as I could saw Malu's father on the road. (1, 127)  
(When I saw Malu's father lying in the middle of the road, I rushed towards him to help.)
- (8) \*He runs no attention to see on his left. . . (3, 28)  
(He did not care to look left.)

These results suggest that the learners have not mastered either rule for the past tense formation or the role of the modal verbs yet. The past tense is still in the early stages of its being mastered, i.e. it is a non-analysed linguistic form. In Table 7.12 we can see that almost all the errors on this particular verb were committed by the least advanced learners (i.e. 3.7% of all errors at level 1, as opposed to 0% at level 2 and 0.7% at level 3). Most non-errors in relation to this verb involve infinitival complementation of the type:



- (9) \*The car of polishman came to see what's happened. (1, 133)  
 (The Police came to see what had happened.)

In addition, only those at level 3 were able to produce correctly the rather complex construction of the following type:

- (10) I entered in all shops of Galerie 24 November without seeing a bag which could please me. (3, 29)

Therefore, the fact that we do not find errors such as *\*I am seeing a lorry* but, instead, many attempts of AUX *can, could + see* indicates that the learners do grasp the semantics of *see* as a state verb, but are still struggling with its syntactic properties in connection with the irregular past.

### Take

Most errors in relation to this verb type involve the irregular past tense. Here are some of the typical examples found in the corpus:

- (11) \*I was very tired and take a small rest. (3, 1)  
 (I was very tired and took a small rest.)
- (12) \*He went to her bedroom, he take her bicycle. (1, 112)  
 (He went to his bedroom and took his bicycle.)
- (13) \*When Bola taking her bicycle it was 7. 45. (1, 112)  
 (When Bola took his bicycle it was 7. 45.)
- (14) \*He. . . take the direction of the school.  
 He was riding very quickly.  
 (He took the direction of the school; he was riding very fast.)

Among the well-formed instances of the past (i.e. *took*) we found many constructions which could hardly be regarded as appropriate on a semantic level (especially lexical choice), and so, which would probably be considered borderline - grammatically speaking.

- (15) \*The securistes dismounted, took me to bed. . .

(The first-aid workers came down and put me on a stretcher.)

- (16) \*After the work he took and got on his bicycle. (1, 144)  
 (After work, he took his bicycle and got on it.)

This last example (16), seems to run against Gass (1984) and Schmidt's (1980) suggestion that the types of co-ordination used by L<sub>2</sub> learners followed universal constraints. In that study, Schmidt examined such co-ordination phenomena as represented in the following sentences:

(17) \*John plays the violin and Mary the piano.

(18) \*John the violin and Mary plays the piano.

Schmidt found that her subjects who came from a wide-range of L<sub>2</sub> backgrounds did not regard sentence (18) as acceptable. Owing to processing difficulties, however, learners may well produce inaccurate utterances, which given more time they may well judge as non-grammatical.

### Have

Most errors committed in relation to *have* involve subject-verb agreement, and the past tense.

- (19) \*He never has ever his dinner in that morning. (1, 114)  
 (He didn't have his breakfast that morning.)

- (20) \*It is Remond the man who have her bicycle. (1, 117)  
 (This story is about a cyclist called Remond.)

- (21) \*At the Boulevard 30 Juin we have a traffic-jam. (1, 135)  
 (There is a traffic-jam in the Boulevard 30 Juin.)

In addition, *have* is being used in a rather vague way as a substitute for a wide range of more specific verbs or nouns as in \**we have a traffic-jam* instead of *there is . . .* or \**the man who have her bicycle* (i.e. a cyclist).

The results in Table 7.12, however, seem to suggest that *have* may involve variable strategic importance at different levels. On the one hand, the verb *have* is used by those learners both at the most advanced and the lowest levels (1.2% at level 3, that is, twice as often as at level 1 with only 0.6%). On the other hand, the use of *have* as a relational verb (Quirk et al. 1972: 96) has been completely avoided at level 2. There is little doubt that these learners can decode this particular meaning of *have*. Just why they were unwilling to use this form is not clear. It may well be that they are not so sure whether this kind of paraphrase would enable them to say exactly what they mean; thus, they avoided taking 'low-gain high-costs risks' (Beebe 1983) altogether, and so tried to use a more precise term to eliminate the danger of being misunderstood.<sup>2</sup>

### 7.9.2. Intransitive Verbs

This category include a variety of verbs most of which fall into the semantic sub-class of so-called 'action' verbs or 'event' verbs (Leech 1971: 18), especially transitional event verbs (e.g. *arrive, come, fall, stop, etc.*), momentary event verbs (e.g. *collide*), and also activity verbs (e.g. *run, walk*). Verbs such as these share one characteristic in common: they take the progressive form *-ing*, i.e. the so-called 'expanded form' (Zydattiss, 1976) as opposed to the simple form of the verb. Not surprisingly, the main learning problem appears to be in relation to the progressive aspect. However, since aspect is so intricately linked with tense, the latter has also been causing some problems to the learner's performance. Let us examine each of these verbs in detail.

---

<sup>2</sup>Beebe (1983) gives a comprehensive review of most of the recent literature concerning learner language strategies, especially 'risk-taking' and 'avoidance'.

## Arrive

Errors in relation to this verb involve mainly the concord between subject and verb in 3rd person singular, both for simple present and progressive present.

- (22) \*When he arrive late the headmaster punishes her. (1, 112)  
(Whenever he arrives late, the headmaster punishes him.)
- (23) \*There is a bicycle man who arrive, but I don't see this bicycle man.  
(A cyclist was coming but I (the lorry driver) couldn't see him.)
- (24) \*Sudnely, he arrives to the car, and he don't stop. (1, 105)  
(Suddenly he found himself too close to the car but he couldn't stop.)

Looking at Table 7.13, we can see that all attempts made by learners at level 2 were successful, whereas at level 1, the frequency of errors was 1.8% and at level 3, we found 1.3% of all errors. Among the non-errors, though, sometimes the learners appear to use the form of the verb correctly (in terms of past tense, negation, etc.). But the problem often arises in connection with the choice of the right preposition to actually represent the intended meaning, as in (24) above.

## Come

Aspectual distinction rather than tense accounted for most of the errors involving this verb type.

- (25) \*This person came from the work. (1, 158)  
(That person was coming from work.)
- (26) \*He has a bicycle that he comes everydays in Kasi Institute. (1, 108)  
(He comes to Kasai Institute everyday on his bicycle.)
- (27) \*He didn't see at time a car which coming in the perdendicular sens. (2, 94)  
(He didn't see in time a car which was coming along in the perpendicular direction.)

Table 7.13

## Verb Phrase Performance Features: Intransitive Verbs

Features	Level 1			Level 2			Level 3			Total		
	+	*	A	+	*	A	+	*	A	+	*	A
<i>arrive</i>	n= 37 %= 1.8	6 1.8	43 1.8	7 1.6	0 0	7 1.5	15 1.2	2 1.3	17 1.2	59 1.6	8 2.9	67 1.6
<i>come</i>	n= 34 %= 1.7	8 2.4	42 1.8	7 1.6	1 2.1	8 1.7	20 1.5	6 4.1	26 1.8	61 1.6	15 2.9	76 1.8
<i>collide</i>	n= 1 %= 0	13 4.0	14 0.6	1 0.2	2 4.2	3 0.6	0 0	0 0	0 0	2 0.1	15 2.9	17 0.4
<i>fall down</i>	n= 11 %= 0.5	5 1.5	16 0.7	0 0	1 2.1	1 0.2	3 0.2	1 0.7	4 0.3	14 0.4	7 1.3	21 0.5
<i>get</i>	n= 13 %= 0.6	11 3.4	24 1.0	2 0.5	0 0	2 0.4	10 0.8	2 1.3	12 0.8	25 0.7	13 2.5	38 0.9
<i>happen</i>	n= 8 %= 0.4	2 0.6	10 0.4	2 0.5	0 0	2 0.4	4 0.3	3 2.0	7 0.5	14 0.4	5 1.0	19 0.4
<i>live</i>	n= 13 %= 0.6	5 1.5	18 0.8	1 0.2	0 0	1 0.2	1 0.1	2 1.3	3 0.2	15 0.4	7 1.3	22 0.5
<i>pass</i>	n= 4 %= 0.2	7 2.1	11 0.5	3 0.7	0 0	3 0.6	9 0.7	1 0.7	10 0.7	16 0.4	8 1.5	24 0.6
<i>run</i>	n= 4 %= 0.2	5 1.5	9 0.4	0 0	0 0	0 0	4 0.3	3 2.0	7 0.5	8 0.2	8 1.5	16 0.4
<i>stop</i>	n= 12 %= 0.6	4 1.2	16 0.7	0 0	0 0	0 0	14 1.1	2 1.3	16 1.1	26 0.7	6 1.1	32 0.7
<i>walk</i>	n= 1 %= 0	4 1.2	5 0.2	0 0	0 0	0 0	6 0.5	2 1.3	8 0.6	7 0.2	6 1.1	13 0.3
Total Intransitive Verbs												
n=	138	70	208	23	4	27	86	24	110	247	98	345
%=	6.7	21.4	8.7	5.3	8.3	5.6	6.7	16.1	7.6	6.5	18.7	8.0

Key:

- + = correct form
- \* = incorrect use
- A = frequency of attempts to use form

Looking at these results (e.g. (25) above) we can argue that in many instances aspect is the learning problem. In example (25), the simple past irregular is correct, but considering the context, we realize that it is the progressive form which is required. In sentence (26), there does not seem to be any morphological errors either, but the context in which this construction is made renders it somewhat incoherent and grammatically ill-formed. Such attempts (as well as (27)) deserve great care on the tutor's behalf, i.e. the latter could advise the learner on how to re-organise information in context to make the utterance more coherent.

### Collide

Many errors involving this verb are due to the wrong allocation of *collide* to the sub-category of transitive verbs; or sometimes to the wrong choice of prepositions.

(28) \*Soudanly, the man who ride the bicycle collide to the car. (1, 163)  
(Suddenly the cyclist collided with the lorry.)

(29) \*It was impossible to brake. He collide against the car. (1, 145)  
(It was impossible to brake. He collided with the car.)

Other errors seem to involve the regular past as in example (29). But they are less severe, since, presumably, the learners in this case were concentrating more on the meaning of the sentence, thus, they overlooked the morphological marking of the verb. These results also show that (in Table 7.13) learners at level 3 seem to be avoiding the verb *collide*; instead, they used forms such as *knocked the car* which roughly conveys the same idea.

### Fall

Transitional event verbs like *fall* normally denote transition into a state rather than the state itself. In this sense, *fall* would be used with progressive to indicate an approach to the transition. However, these results suggest that *fall* is being used most of the time as a momentary verb (i.e. like *jump*, *hit*, etc.) to indicate, as it were, a happening so momentary that it does not seem to have any duration. In the following examples, no progressive form was used and sometimes a verb such as *drop* would have been a better choice.

- (30) \*When I hurt me, I fall down with my bike. (3, 43)  
 (After hurting myself, I fell with my bike.)
- (31) \*He feel down on his bicycle. (1, 107)  
 (He fell off his bicycle.)

Most errors here are thus due to incorrect use of irregular past, and sometimes the choice of prepositions to go with the verb.

### Get

Most errors found here are in connection with the irregular past (i.e. *got*). Similarly, as we expected, the choice of prepositions constitute another difficulty since the latter occupy a borderline zone between the syntax and semantics of English.

- (32) \*I stopped and get out of the lorry. (3, 15)  
 (I stopped and got out of the lorry.)
- (33) \*One day Remond get of his home with angry because he is late  
 for to work.  
 (One day Remond left his home angry because he was late.)

Looking at the performance of the different levels of learners, we realize that level 2 performed better since all attempts of *get* were correctly used, i.e. 0% errors, whereas level 1 committed errors up to 3.4% of all errors and level 3, only 1.3%

## Happen

Most errors involving *happen* are related to both the present and past perfect aspect. For example,

- (34) \*I go to tell you what happening for a young man. (1, 119)  
(I'm going to tell you what happened to a young man.)
- (35) \*The man is comming to se what happen. (3, 76)  
(The man is coming to see what has just happened.)
- (36) \*A passenger who came to see to what happened signaled to the hospital. (1, 148)  
(A passer-by contacted the hospital after seeing what had just happened.)

Again, learners at level 2 were performing better since all of their attempts were correct. The other two levels made equal number of attempts (0.4% of the total each), with level 1 committing errors less frequently than level 3.

## Live

Together with *happen*, *live*, was the only state verb as opposed to action verbs. One of the main characteristics of state verbs is that they do not normally take the progressive form (i.e. - progressive, -imperative (Lakoff 1966; Chafe 1970)). This requirement was not violated by these learners. Of course, this absence of a problem or error does not necessarily mean that the way *live* as a state verb functions in English, is fully understood. It only means, under the present circumstances, that the learners responded in a target-like manner. Let us look at the most typical examples:

- (37) \*He left in a small house near the big market. (3, 56)  
(He lived in a small house near the big market.)
- (38) \*This young man is Dick Smith. He live in a big city. (1, 120)  
(This young man is Dick Smith. He lives in a big city.)



From these results, it appears that many errors involving *live* were due to the syntax of English (i.e. the marking of the 3rd person singular) rather than semantic considerations. In addition, in example (37), there was a confusion at the phonological level between two verbs (i.e. *live* and *leave*), leading to semantic error. This error cause may be attributed to overgeneralization of the correct irregular past of *leave* (i.e. *left*) to a different verb type, the regular verb *live*.

### Pass

The main problem in relation to this verb is the use of the progressive aspect. This verb falls into the category of process verbs (Leech 1971: 19). Such verbs are normally compatible with the progressive aspect, since they ordinarily involve duration, as in, e.g. *The soup is thickening*; but this duration is not infinite duration. Let us look at some of the examples:

- (39) \*He saw a bag car which passed. . . in the man the road. (1, 138)  
(He saw a big lorry which was passing along the main road.)
- (40) \* . . . he has seen his master passed with his car. (3, 10)  
(. . .he saw his boss pass in his car.)
- (41) \* . . . it was a great lorry that passed. (3, 10)  
(A big lorry was passing along the street.)

The results in Table 7.13 again show the learners at level 2 performing better (0% errors) than level 1 (2.1%) and level 3 (0.7%).

### Run

Almost all the errors committed in this area were related to the irregular past. Although the simple past was required, in many contexts the form *ran* was not used by the respondents. Clearly, these learners were unaware of its existence;

probably they assumed *run* does not change to indicate simple past. Here are some of the examples:

- (42) \*I jumped on my bicycle and I was in speed.  
I run away. (3, 24)  
(I jumped on my bicycle and ran away quickly.)
- (43) \*Remond run after her in the boulevard. (1, 117)  
(Remond ran after her. . .)
- (44) \*Bankaka running for his father to tell him  
the news. (1, 158)  
(Bankaka ran away to tell his father the news.)

When we look at the non-errors, we realize that they are likely to occur only when *run* is the main predicate in relation to an infinitival subordinate clause, as in e.g. *So I couldn't run very fast.* (3, 43), or *He began to run quickly to arrive before the night* (1, 130). This tendency, as we see, was established for both levels 1 and 3. The results in Table 7.13 show, on the other hand, that at level 2, this verb was completely absent. In other words, while the other two levels (1 and 3) tended to generalize the form *run* to both simple present and past (probably equating *run* with such verbs as *cut*, *hit*, etc.), respondents at level 2 were using the strategy of avoidance to cope with this difficulty.

### Stop

Errors involving this verb show that it is not only the irregular past which was still causing problems but also the regular past.

- (45) \*The police stop he and they came. (1, 101)  
(The police stopped him and came over.)
- (46) \*The driver stop his car and saw the rider. (1, 161)  
(The driver stopped the car and saw the cyclist.)
- (47) \*I do many efforts to go away but the car can't stopped. (3, 4)  
(I tried hard to give way to the lorry but it didn't stop.)

Apart from the failure to mark the verb for simple past (e.g. (45) and (46)), we also notice, as in e.g. (47) that some of the more advanced learners have reached a pre-systematic stage in relation to the regular past. In other words, some of the advanced students (i.e. level 3) were able to mark correctly the verb for regular past, but they often ran into difficulty when the verb was used with the auxiliaries (see (47) *can't stopped*). Again, the strategy of avoidance seems to have been the solution adopted at level 2 to deal with this problem.

### Walk

Errors committed in relation to this verb derive from three main sources: 1) phonological confusion, 2) progressive aspect, and 3) lexis. For example,

(48) \*In the Monday morning he walk up only late. (1, 114)  
(On Monday morning, he woke up late.)

(49) \*One day Wenceslas decided to walk with his bicycle.  
(1, 145)  
(One day Wenceslas decided to travel by bicycle.)

(50) \*I dreamed that when I have walking I have done an  
accident. (3, 34)  
(I dreamt that while I was walking I met with an accident.)

On the phonological level, there is a confusion between *walk* and *wake up*, leading to incorrect use, in e.g. (48) above. On the syntactic level, there is the more frequent problem of distinguishing between the simple form and expanded form (or progressive aspect), as in e.g. \**The man couldn't carry his bicycle because he walked very fast.* i.e. *The man couldn't control his bicycle because he was riding fast.*

On the lexical level, the verb *walk* was often used vaguely or wrongly to express various idiosyncratic meanings. These included *to walk* but also *to travel*, *to ride*, which all involve the idea of using a particular means of

transport. The only L<sub>2</sub> approximation (Drubig 1972; Zydattiss 1976) which resulted in real trouble (or error) was that found in e.g. (48) above where *I walk up* is produced instead of *I woke up* because of the phonological confusion, probably traceable to inadequate teaching techniques. These results suggest that, in general, with intransitive verbs, overgeneralization is the main strategy being resorted to more frequently. Little evidence, if any, of transfer from French or L<sub>1</sub> is apparent. However, learners at the university (i.e. level 2) make extensive use of the strategy of avoidance. In this connection, at level 2, four verbs out of eleven in relation to the intransitive verbs were not attempted at all; that is, 36% of these forms were avoided altogether at level 2, as opposed to 0% at level 1 and 9% at level 3.

### 7.9.3. Copula *Be*

Concord was the main source of difficulty with *be* copula (i.e. as an equational or existential verb rather than an auxiliary). For example,

- (51) \*But at the hospital, no ambulance were there. (3, 9)  
(But no ambulance was available at the hospital.)
- (52) \*When the young man were sick, the officials transported him to the hospital. (3, 57)  
(But no ambulance was available at the hospital at the time.)
- (53) \*There were a big car which just appear. (2, 97)  
(There was a big lorry coming from the right-hand side of the street.)

These results (in Table 7.14) suggest that *be* is not so frequently used by these respondents (0.5% of all attempts). However, its accuracy rate is high (15 out of 20 attempts are non-errors, i.e. 75%). Respondents at level 1 scored 9 non-errors out of 9 attempts (i.e. 100%), whereas level 2 attempted to use *be* only once and failed. At level 3 the accuracy rate was 60%. Overall, the learners were having difficulty with the third person singular for the simple past. They

were using consistently, the form *were* (which is the form for plural or the 2nd person). Although they were producing correct utterances including e.g. *it wasn't the case* (3, 43), or *it wasn't my fault. . .* (3, 28), these phrases could not be taken as evidence for concluding that the learners had mastered the copula. In fact, they used *wasn't* only with the negative form but used *were* + 1st or 3rd person singular elsewhere.

#### 7.9.4. Auxiliaries

##### Have

Errors related to *have* as an auxiliary verb involve the concordance of tenses as well as aspect. For example,

- (54) \*In spite of he has seen the lorry near a square he continued to ride fast his bicycle. (1, 152)  
(He continued to ride his bicycle even though he had seen the lorry near the square.)
- (55) \*What did happen when I have reach the level of theis street, place where an accident have took place. (3, 64)  
(You may be wondering what happened when I reached the street level where the accident had taken place.)

These results exemplify the kind of difficulty that the learners were running into. First, there is the concordance of tenses between the main and subordinate clauses. There is also the combination of past tense formation and asking questions about two predicates representing actions or state of affairs situated at two different points in time.

Overall, respondents at the level 2 were doing better than both level 1 and 3. At level 2, the only one attempt made was successful. However, at level 1, all six attempts resulted in errors, whereas at level 3, ten out of thirteen were non-errors (i.e. 77% accuracy rate). This particular structure has to be dealt

Table 7.14

## Overall Verb Phrase Performance: All features

Features	Level 1			Level 2			Level 3			Total		
	+	*	A	+	*	A	+	*	A	+	*	A
<b>1. Transitive Verbs</b>												
n=	122	33	155	21	10	31	91	16	107	234	59	293
%=	5.9	10.1	6.5	4.8	20.8	6.4	7.0	10.7	7.4	6.2	11.3	6.8
<b>2. Intransitive Verbs</b>												
n=	138	70	208	23	4	27	86	24	110	247	98	345
%=	6.7	21.4	8.7	5.3	8.3	5.6	6.7	16.1	7.6	6.5	18.7	8.0
<b>3. BE as copula</b>												
n=	9	0	9	0	1	1	6	4	10	15	5	20
%=	0.4	0	0.4	0	2.1	0.2	0.5	2.7	0.7	0.4	1.0	0.5
<b>4. Auxiliaries</b>												
<i>Have</i>												
n=	0	6	6	1	0	1	10	3	13	11	9	20
%=	0	1.6	0.3	0.2	0	0.2	0.8	2.0	0.9	0.3	1.7	0.5
<i>Do</i>												
n=	25	15	40	8	1	9	28	5	33	61	21	82
%=	1.2	4.6	1.7	1.8	2.1	1.9	2.2	3.4	2.3	1.6	4.0	1.9
<b>5. Subordination</b>												
<i>When / while</i>												
n=	60	2	62	11	1	12	35	5	40	106	8	114
%=	2.9	0.6	2.6	2.5	2.1	1.9	2.7	3.4	2.8	2.8	1.5	2.6
<i>in order to</i>												
n=	0	11	11	0	1	1	2	0	2	2	12	14
%=	0	3.4	0.5	0	2.1	0.2	0.2	0.0	0.1	0.1	2.3	0.3
<b>Total for Verb Phrase</b>												
n=	354	137	491	64	18	82	258	57	315	676	212	888
%=	17.2	41.9	20.6	14.8	37.5	17.0	20.0	38.3	21.9	17.9	40.5	20.6

Key:

- + = correct form
- \* = incorrect use
- A = frequency of attempts

with more systematically by teachers, especially at level 1.

### Do – support

Many errors involving *do support* are due to the fact that learners use this form as a substitute for *can* or to indicate negation but not tense. For example,

- (56) \*I don't listen very well but I saw the policeman  
telephone to the hospital. (3, 16)  
(I couldn't hear very well, but I saw the policeman making  
a phone call to get an ambulance.)
- (57) \*He don't heard when the car was giving  
the warning. (2, 96)  
(He didn't hear the warning given by the car.)
- (58) \*But behind her, Nr. Bola don look. (1, 133)  
(But Mr. Bola didn't look ahead of him.)

These results (Table 7.14) indicate *do support* was used more frequently than *have* (i.e. four times as often as *have*) and that the overall performance at all 3 levels was better too (i.e. 74% accuracy rate for *do* as opposed to 55% for *have*). We find again that level 2 was making errors less frequently (1 error out of 9 attempts, i.e. 11%) than level 1 (15 errors out of 45, i.e. 33%) and level 3 (5 out of 33, i.e. 15%). The confusion between the meaning of *do* and *can* is apparent in example (56) above where *don't listen* is produced instead of *couldn't hear*. Examples (57) and (58), on the other hand, suggest that the auxiliary known as *do support* is still in a pre-systematic stage, or has not undergone adequate linguistic analysis. This is, perhaps, a learner-external error cause which could be remedied through an appropriate pedagogic grammar.

### 7.9.5. Subordination

The part of the learner's performance subsumed under this sub-category includes 'temporal frame' and 'infinitival complementation'. Errors in relation to subordination represent 3.8% of all errors when all three levels are combined.

In the following examples, we will look closely at both features found within this sub-category.

**Temporal Frame: while /when**

The notion of temporal frame follows from the notion of limited duration in connection with the progressive aspect. According to Leech (1971: 17), whenever a point of time or event is in a contemporaneous relation with a happening of duration, it is natural that the durational happening should overlap the durationless event or point in both directions – in short, a temporal frame should be set up. Errors with the 'Inzidenzschema' (Pollak 1960: 129) reflect the learner's unsuccessful attempt to convey this framing effect as regards the use of one of the two following forms: *when* or *while*. For example,

(59) \*When he rode he didn't pay attention in front of him. (2, 77)  
(While riding he didn't look ahead.)

(60) \*When he was riding he didn't pay attention. (1, 145)  
(While he was riding he didn't pay attention.)

(61) \*When I did my travel, I didn't hear noise of engine. (3, 43)  
(During my journey, I didn't hear the sound of the hooter.)

Looking at these results we realize that *when* is often substituted for *while*, and this leads to an error of incidence. At level 2, we did not find an error involving the form *while*. But this does not mean that these learners have mastered the semantics of temporal frame. In fact, in example (59), we notice that they produced *when* in a context where they should have used *while*. At all levels this confusion is apparent. Nehls (1974: 84) has suggested that 'Incidence' can be realized not only within the complex sentence but also beyond the sentence boundary, as in (61).



### 7.9.6. Infinitival Complementation

Errors in relation to this sub-category are due to the use of the phrase *\*for to* instead of *in order to* to introduce a purpose subordinate clause. Other construction types involving infinitival complements ( or nominalizations) were dealt with in detail in one of the preceding chapters and will not concern us here.<sup>3</sup> For example,

- (62) \*His father went with Jim for to buy a bicycle  
for him. (1, 124)  
(Jim's father went with him to buy him a bicycle.)
- (63) \*He put the new bicycle for to go at work. (1, 117)  
(He used the new bicycle to go to work.)
- (64) \*He woke up at five past seven, so fifty minutes left  
for to reach the school. (2, 91)  
(He woke up at five past seven, so he had fifty minutes  
to reach his school.)

These results (in Table 7.14) indicate that almost all of the errors were committed at level 1 (11 out of 12), and none was found at level 3. All the attempts made at level 3 were successful. They (level 3) used the grammatically correct phrase *in order to* rather than *for to*. As usual, respondents at level 2 were rather cautious and attempted to use *for to* only once.

### 7.10. Learner Language Behaviour: Prepositional Phrase Related Features

Several studies have pointed to the difficulties encountered by L<sub>2</sub> learners of English as regards prepositions and prepositional phrases. The major cause of these difficulties is that the prepositional phrase (PP) is in many ways at the interface of syntax and meaning. In our analysis of the learner's performance

---

<sup>3</sup>See Chapter 6 on *Promise, ask, tell*

both the syntactic function of the PP as well as its semantics will be considered.

Quirk et al. (1972: 299) define a prepositional phrase as 'a preposition followed by a prepositional complement, which is characteristically either a noun phrase or a clause (wh-clause or V-ing clause) in nominal function'. e. g.:

(Preposition) *with* - gratitude  
                   *in* - the garden  
                   *from* - what we know  
                   *by* - offering some help

From a syntactic viewpoint, prepositional phrases may function as:

- a.- Adjunct: e.g. I jumped on the river bank.
- b.- Postmodifier: e.g. They visited a village in the borders.
- c.- Complementation of a verb: She looked at her picture and smiled.
- d.- Complementation of an adjective: Tony was happy for his sister.
- e.- Disjunct: They still love each other, in spite of all their bad fortunes.
- f.- Conjunct: Of course, he will come if he can.

Among these functions, (a), (b) and (c) appeared more frequently within the corpus and so they will concern us most.

From the semantic point of view, a preposition expresses a relation between two entities, one being that represented by the prepositional complement (Quirk et al. op. cit.). There are various types of relational meaning, but those of PLACE and TIME are the most prominent. Other relationships include: INSTRUMENT, CAUSE, MEANS, ACCOMPANIMENT, POSTMODIFIER OF NOUN PHRASE, etc. In our analysis we will deal with prepositional phrases by allocating them to one of the following meaning types:

- 1.) Place (i.e. *to, at, in, on*)
- 2.) Time (i.e. *at, on, in*)
- 3.) Post-modifier of noun phrase (i.e. *of*)
- 4.) Other Adjuncts (i.e. *for, with, by*)

### 7.10.1. Learner's Performance on the Prepositional Phrase

#### 7.10.1.1. Place

##### To

Prepositional Phrases (PP) of place used with *to* typically accompany action verbs of dynamic 'motional' meaning (e.g. *go, come, etc.*). The errors found in these data are mainly due to the use of *to* with PPs of position, which are normally associated with stative verbs. e. g.

- (1) He was very late to his work.  
(It was very late when he arrived at work.)
- (2) So he cracked to a lorry. (1, 131)  
(So he crashed against the lorry.)
- (3) I passed to the town for to see some thinks. (1, 156)  
(I was passing in town looking around in the shops.)

These results suggest that the meaning of *to* was more frequently misconstrued by learners at level 1 (see Table 7.15), than those at the other levels. No errors in connection with this meaning type were found at level 2 or 3. Some kind of  $L_1$  interference of French seems to underlie these learners behaviour.

##### At

Generally speaking, a PP of position can accompany any verb, although this meaning is particularly associated with verbs of stative meaning (e.g. *be, stand,*

Table 7.15

## Prepositional Phrase Performance Features

Features	Level 1			Level 2			Level 3			Total			
	+	*	A	+	*	A	+	*	A	+	*	A	
<i>To</i>	n=	219	17	236	39	3	42	132	7	139	390	27	417
	%=	10.9	5.2	9.9	9.0	6.3	8.7	10.2	4.7	9.6	10.3	5.2	9.7
<i>In</i>	n=	121	26	147	32	4	36	60	14	74	213	44	257
	%=	5.9	8.0	6.2	7.4	8.3	7.5	4.6	9.4	5.1	5.6	8.4	6.0
<i>Of</i>	n=	53	6	59	17	0	17	56	4	60	126	10	136
	%=	2.6	1.8	2.5	3.9	0	3.5	4.3	2.7	4.2	3.3	1.9	3.2
<i>At</i>	n=	62	19	81	9	2	11	31	4	35	102	25	127
	%=	3.0	5.8	3.4	2.1	4.2	2.3	2.4	2.7	2.4	2.7	4.8	2.9
<i>On</i>	n=	44	6	50	11	1	12	46	5	51	101	12	113
	%=	2.1	1.8	2.1	2.5	2.1	2.5	3.6	3.4	3.5	2.7	2.3	2.6
<i>For</i>	n=	23	9	33	3	1	4	24	1	25	50	11	61
	%=	1.1	2.8	1.3	0.7	2.1	0.8	1.9	0.7	1.7	1.3	2.1	1.4
<i>With</i>	n=	21	4	25	5	0	5	15	2	17	41	6	47
	%=	1.0	1.2	1.0	1.2	0	1.0	1.2	1.3	1.2	1.1	1.1	1.1
<i>By</i>	n=	21	3	24	5	0	5	12	4	16	38	7	45
	%=	1.0	0.9	1.0	1.2	0	1.0	0.9	2.7	1.1	1.0	1.3	1.0
Total For Prepositional Phrase													
	n=	564	90	654	121	11	132	376	41	417	1061	142	1203
	%=	27.4	27.5	27.4	27.9	22.9	27.4	29.1	27.5	28.9	28.0	27.1	27.9

Key:

- + = correct form
- \* = incorrect use
- A = frequency of attempts

*live*). *At* indicates a dimensionless location, e.g. a point on the map with no details regarding its shape or size.

- (4) \*Mr. X wants to go shopping at the London city. (3, 74)  
(Mr. X wanted to go to London to do some shopping.)
- (5) \*This street it was at my left. (3, 15)  
(This street was on my left.)
- (6) \*At Londres, everybody prefers to rest on the weekends. (1, 138)  
(In London, everybody enjoys taking some rest at the weekend.)

Instead of being used with stative verbs to convey dimensionless positional meaning, *at* was being used with action verbs (e.g. *go*) expressing destination meaning or with two or three dimensional objects (e.g. \**look at my left, at Londres*) in (5) and (6) above.

## On

Some of the errors involving this preposition are due to its being used with dynamic verbs involving the meaning of destination. Other errors derive from overgeneralization (of *on*) to contexts where 0 preposition would be required in English.

- (7) I was riding on bicycle in the morning at 7: 30. (3, 20)  
(I was riding my bicycle in the morning at 7: 30.)
- (8) A lorry come out from opposite side on such high speed. (3, 60)  
(A lorry came from the opposite direction at high speed.)
- (9) He fall down on a big puddle on the road. (1, 134)  
(He fell into a big puddle on the road.)
- (10) He crashed the car and fell down on his bicycle. (1, 107)  
(He crashed into the lorry and fell from his bicycle.)

In example (7), *on* is redundant since the semantics of *ride* imply that the human agent must be positioned *on* the bicycle. In addition, in (8), *on* is used as an adjunct (instead of *at*) to relate the dynamic verb *come* with a manner

phrase *high speed*.

Looking at the frequency of non-errors however, we can see (Table 7.15) that 101 out of 113 total attempts were correct, i.e. 89%. At level 1, the accuracy rate was 88% compared with 92% at level 2 and 90% at level 3.

### In

Errors involving *in* are due to the fact that its function is still in a pre-systematic stage. This preposition is normally used to indicate PPs of place which have three dimensions, i.e. an area, or volume. However, in the following examples, *in* is not used in this sense.

- (11) \*I was thrown in the other side from my bicycle. (3, 20)  
(I was thrown away on the other side.)
- (12) \*So he has not paid attention in what would be happened. (3, 51)  
(. . .he did not pay attention to what might happen.)
- (13) \*He wanted to go in the centre of town. (2, 77)  
(He wanted to go to the city centre.)
- (14) \*There was no brakes in the bicycle. (1, 145)  
(There were no brakes on the bicycle or, The bicycle had no brakes.)
- (15) \*When the man arrived in the place of croisement. (1, 125)  
(When the man arrived at the round about. . .)

These results indicate that *in* was often used to relate a verb with a one or two-dimension object (i.e. a point or a surface) instead of a three-dimension object (i.e. an area) as in (11), (14), and (15) above. Sometimes *in* is used wrongly to indicate an adjunct to the verb (e.g. (12)), or with verbs of dynamic motional meaning (e.g. (13) above). In addition, Table 7.15 shows that *in* has the greatest relative frequency of errors (i.e. 8.4% of all errors, involved this type of preposition). Transfer from French appears to be the main cause of the errors, since the French equivalent of *in* (i.e. *dans*) may convey a wide range of

meaning of shades, including those intended by the learners in this area.

#### 7.10.1.2. Time

One important feature of prepositional meaning as regards temporal adjuncts is that sometimes the preposition of time is absent. In this area, the temporal adjunct takes the form of a noun phrase rather than a prepositional phrase, as in

(16) He went to the Elikya Institute last month.

Constructions such as these (i.e. with 0 preposition of time) did not appear to be a source of difficulty for our respondents. Thus, they were not specifically examined. On the other hand, some of the prepositional referred to earlier as preposition of place will be re-examined under this heading (time), depending on the meaning types which they are assumed to express within a particular context.

#### At

The main difference between preposition of place and time is perhaps the fact that in the time sphere, there are only two 'dimension-types', viz. 'point of time' and 'period of time' (place has three dimension types involving a line, surface and volume).

(17) He must be quickly to arrive at time. (1, 115)  
(He must be quick to arrive on time.)

(18) At Londres, everybody preferes to reste on the week-end. (1, 138)  
(In London, everybody enjoys having some rest at the weekend.)

These results show that errors in relation to *at* as a point of time, were committed only at levels 1 and 2. Contrary to normal use, *at* was sometimes used by these learners interchangeably with *on* (e.g. in (17)) or *in* as in (18)

above. These errors are clearly due to interference (viz. the French equivalent in both cases is 'à temps'. Other errors (as in (17)) are due to the idiomatic use of *at* in English for period of time (e.g. at Easter and at the weekend) instead of *on*, which is used to refer to days (e.g. *On Saturdays and Sundays, but at weekends*).

### On

Errors in relation to *on* as a preposition of time were rather infrequent and those that we identified were committed by level 1 respondents only. In the outgoing paragraph (see *at* as a preposition of time) relevant comments are made on these errors. There were 12 non-errors in the sample as regards *on*, as a preposition of time out of a total of 101 correct uses of this form (i.e. 12% of correct uses involved *on* as temporal adjunct). No evidence of the avoidance strategy was found, since the correct uses were fairly well distributed between the three different levels (i.e. 6 attempts were found at level 1, 2 attempts at level 2, and 4 attempts at level 3).

### In

The temporal use of *in* refers to periods of time (e.g. *in the morning, in January, in 1979*). Errors committed in relation to this prepositional meaning involve its generalization to other dimension-types (especially *on* as point of time). For example,

(19) And in the Monday morning he walk up late. (1, 114)  
(And on Monday morning he woke up late.)

(20) And quickly the ambulance arrived in the same time then  
some policemen. (1, 162)  
(An the ambulance arrived quickly, at the same time as  
some policemen.)



Again, these results show that all the errors involving *in* were committed at level 1. Out of the 213 total number of correct uses of *in*, 12 were found at level 1 (0.6%), 3 at level 2 (0.1%) and 5 at level 3 (0.2%).

#### 7.10.1.3. Post-modifier of noun phrase *of*

The most common English preposition *of* occurs mainly as a postmodifying genitive in noun phrases. In this sample, *of* came in 3rd position after *to* and *in* (3.2% of all attempted uses). Some of the errors are due to a phonological problem, viz. the confusion between *of* and *off*. For example:

(21) \*He get of home at 6. 15. (3, 43)  
 (He was off at 6.15) or  
 (He left home at 6. 15.)

(22) \*He lived far of the town. (1, 143)  
 (He lived far off the city centre.)

(23) \*I entered in all shop of galerie du 24 Novembre. (3, 29)  
 (I went into all the shops in the Galerie du 24 Novembre.)

Other errors involved the use of *of* instead of *in* in the sense of an area (e.g. (23)). Apart from these few incorrect uses, *of* was generally used in the normal sense (as a genitive in NPs). For instance, at level 2, all 17 out of 17 attempts were correctly used. No avoidance was adopted as a strategy to hide any learning problem in this particular preposition type. Mainly, overgeneralization was the cause of the errors found at both level 1 (6 errors out of 59 attempts, i.e. 10%) and level 3 (4 errors out of 60 attempts, i.e. 7%).

#### 7.10.1.4. Other Adjuncts: *for*, *with*, *by*

##### For

Some of the most commonest uses of *for* in Standard English involve the notion of purpose. They can usually be paraphrased by a clause introduced by

*in order to* (e.g. *He will do everything for money*). Other uses involve intended recipient (i.e. human or animate NPs) as in e.g. *Mummy cooked a very good meal for us*. Most errors found in the data were related to the notion of purpose rather than intended recipient.

- (24) So fifty minutes left for to reach the school. (2, 91)  
(So he had 50 minutes left to reach the school.)
- (25) And for to arrive there, it was again 3 km. (1, 130)  
(In order to reach the school, he had 3 km more to go.)
- (26) He was riding very fast for a small object to pass the city centre. (1, 150)  
(He was riding very fast because of this simple reason: to get through the city.)

We have already described this learning problem elsewhere, which is due to the ambivalence of the syntax (form) and semantics (function) of the infinitival complements. Here we will only point out that many errors regarding *for*, especially at level 1, are due to the learners' attempts to use *for* with its prepositional meaning of purpose. However, when the PP is a subordinate clause rather than a NP, then, learners run into difficulty and produce errors such as those found in (24) and (25) above. The non-errors accounted for about 82% of all the attempted uses of *for*. Level 1 had an accuracy rate of 72% (23 non-errors out of 32 uses); level 2 reached 75%, and level 3 scored a high 96% accuracy level (24 correct out of 25 uses). This suggests a systematic improvement of the performance if these learners as regards this meaning type.

### **By / With**

Either *by* or *with* can be used to express the notion of means, or instrument. *By* is more commonly used to express the meaning *by means of*, as in e.g. *Julian goes to work by bus*. In addition, *with* expresses instrumental meaning e.g.

*I cut the bread with a knife.* Most errors found in the data are related to some confusion between two meaning types (i.e. using *with* in the sense of *means of*):

- (28) I have travelled with my bicycle. (3, 16)  
(I travelled by bicycle.)
- (29) Usually Remond go to work with his bicycle. (1, 117)  
(Usually Remond goes to work by bicycle.)
- (30) I was riding bike in the town. (3, 43)  
(I was riding my bike in the town.)
- (31) He went there by his bicycle. (1, 113)  
(He went there by bicycle.)

These results suggest that the main cause of difficulty is overgeneralization. Some learners are obviously still struggling with such notions as means, instrument, etc. Non-errors accounted for 79 out of 92 attempts as regards both *with* and *by* (i.e. 86% accuracy rate), as opposed to 13 out of 92 errors (i.e. 14%).

### 7.11. Discussion

Our discussion of the results will focus on the two important aspects of the analysis of data, viz. the quantitative as well as the qualitative aspects. The quantitative part will examine the extent to which the variables that we selected enabled us to establish an index of  $L_2$  development which is theoretically viable, methodologically sound and pedagogically practical. The qualitative part will emphasize the significance of learner language behaviour, reassess such concepts as 'error analysis' and investigate the degree of strategic use adopted by the learner when he is faced with a learning problem.

First of all, from the outset, we experienced some difficulty in comparing most of our findings on establishing an  $L_2$  index of development with those of Larsen-Freeman and Strom (1977) and Larsen-Freeman (1983). The main

difficulty in comparing our results with those of these researchers are due to the fact that they used quite different sample sizes (e.g. sometimes 5 subjects per group); and especially, they adopted an interpretation of a T- unit such that they could not obtain the same results if they applied, on the same data, an analytical technique based on information blocks. We feel that the notion of a T-unit is flawed in that, for example, it assumes equal weight for all t-units (i.e. whether the T-unit is composed of only one clause or one main clause and say, five subordinate clauses clustered around the main clause).

Having said that, we experienced some comfort when we found our results did lead to the same outcome at least in one area. We find that, like in Larsen-Freeman (1983) using a picture composition task, the number of errors declines as the level of proficiency increased. This is confirmed by X square (for the errors part) and by ANOVA, but not Scheffe (for the error-free part of the performance). In addition, like in Cook (1979), we find a high correlation between the average clause embedding (ACEW) with the average number of clauses per information-block. This finding establishes a link between the assessment of native speakers written text and L<sub>2</sub> learners. However, correlation between the ACEW and error-free information blocks was very low ( $r = .27$ ,  $p < .002$ ). This suggests that writing complex sentences does not entail improving on accuracy. This fact may be due either to individual characteristics, or simply greater use of the risk-taking/ avoidance strategy.

However, there are indications that, apart from possible individual characteristics, most university students were using avoidance from time to time to deal with particular problems or grammatical features. Nevertheless, as shown by ANOVA, the differences between the levels on error-free information blocks were not big enough to distinguish them (using appropriate Scheffe

tests). Since the ANOVA results deal with the amount of variability in the data rather than frequencies, we are inclined to regard the X square results as indicating only a trend, not a definite pattern, for the university students to adopt the avoidance strategy. A more qualitative analysis of data, i.e. from an error-analysis viewpoint, does however enable us to pinpoint specific learning problems, in particular, regarding the verb phrase, as we will see later in this section.

As regards the usefulness of the ACEW measure, we have been able to establish that this constitutes a highly reliable and valid measure for assessing both the learners' global proficiency and the degree of risk-taking that they were willing to take. We have already mentioned the fact that our results on this account are similar to those found by Cook (1979) applying the same technique of clause analysis. We feel confident that other researchers could easily reduplicate such analyses and obtain comparable results as regards the amount and degree of embedding, whence information processing in their samples; and that the ACEW does differentiate significantly between learners belonging to different levels of training. In connection with risk-taking, we have also argued that the ACEW is a valid measure since it actually indicates how much information the learner or writer is willing to communicate, even though some aspects of it may contain errors. Presumably, those learners who scored low on the ACEW are the ones who were not willing to take or create initiatives to communicate more than what they regard as minimum information required, for fear that they might run into difficulty or make errors. According to the ANOVA and appropriate Scheffe tests results, the more advanced learners were also the ones who significantly, were taking more risks (i.e. they had a bigger ACEW value).

Then there is our proposed notion of communicational capability index (CCI). We originally speculated that, as Larsen-Freeman (1983) suggest, the proportion of error-free information blocks (EFIB) is the favourite candidate for representing  $L_2$  global index. Unfortunately, the EFIB results were not significant by Scheffe tests; and so we decided to rely on the ACEW instead. In order to take into account the distinction between native speaker writers and  $L_2$  learners, we devised the following credit system.

If, for instance, three different subjects (X, Y, Z) have the same value (e.g. 1.61) for Average Clause Embedding Weight (ACEW) but they scored different values (e.g. 0.90, 0.62, and 0.27) for the proportion of error-free blocks (EFIB), then their CCI would be represented using a 3-star (\*\*\*) system as follows: 1.) for subject X, the communicational capability index will be CCI = 1.61 \*\*\*; 2.) for subject Y, the CCI = 1.61\*\*; and 3.) for subject Z, the CCI = 1.61\*. The 3-star credit system is a satisfactory way of assessing global second language learner's potential to encode both elaborate (ACEW) and increasingly comprehensible messages (EFIB) in the target language. This notion of CCI appears to have at least three major advantages compared to earlier attempts to construct indices of  $L_2$  development. First of all, on the theoretical level, the CCI amounts to a model of speech perception and production which attempts to account for different rates of acquisition of perceptual and productive skill without assuming the existence of two separate grammars (i.e. a perception grammar and production grammar). Naiman (1974) and Swain, Dumas and Naiman (1974) maintain that two separate grammars must be assumed to exist in order to account for such distinctly different abilities in the second language learner - a comprehension grammar and a production grammar. However, this view has been rejected by Tarone (1974: 226) who stresses this:

. . . it is my contention that it is unnecessary to assume that two separate linguistic grammars exist, each with its own lexical and syntactic rules.

Referring to both Neisser's (1967) and Bever's (1970) models of speech perception, Tarone (1974) concludes by saying 'as long as researchers are attempting to describe the data in linguistic terms only, they may be binding themselves to patterns and consistencies in the data which are describable in terms of perceptual or productive strategies, and not in terms of linguistic rules.' In addition, the CCI has greater comparability with native speakers texts, in terms of its processing ease or difficulty, since the CCI is, in effect, equivalent to the ACEW plus the 3-star marking system. In other words, the native speaker's average clause embedding weight is unmarked, whereas the L<sub>2</sub> learner is marked, but based on exactly the same assumptions viz. potential for information processing.

Secondly, on the methodological level, an index based on the Information Block Theory such as the CCI has the advantage of being both rater and task independent. In other words, the investigator does not need to rely on a number of independent judges to obtain an objective assessment of L<sub>2</sub> learner's ability in writing; since in fact every judge would come up with exactly the same value if all the steps described in section 7.3 above were followed. Once this index is established, the examiner can save invaluable time and proceed to a more qualitative analysis of the learner's linguistic or discursal processes.

Thirdly, on the pedagogical or practical level, the CCI measure provides classroom practitioners with a reliable and practical means for dealing with written compositions in a convenient way, since it eliminates the judges' subjectivity (even if they could be found on a day-to-day basis in an

educational setting), and since it has communication (or potential for information processing) as its main tenet.

Of course the issue of communication as a criteria for assessing L<sub>2</sub> global proficiency involves not only information processing (i.e. productive and receptive abilities) but also the extent to which the L<sub>2</sub> learner achieves successful communication. This need for a further stage has been fully recognised; but for reasons already given (see section 7.25) no specific attempt to investigate successful communication has been undertaken. What has often appeared controversial, however, is the use of communication as a criterion for error evaluation (Johansson, 1973; Enkvist, 1973; Hughes and Lascaratou, 1982; Davies, 1985). It has been suggested that in the evaluation of learner's errors more use should be made of the criterion of intelligibility. Johansson (1973), for instance, proposes a system of evaluation in which prominence is given to the error's effect on comprehensibility; although he also gives some consideration to other factors such as degree of irritation caused by an error, its frequency, and generality. Enkvist (1973) goes even further and implies that comprehensibility is the only valid criterion and that errors should be equated with breakdown in communication. In addition, Hughes and Lascaratou (1982) deplore the fact that according to their interpretation of some experimental results, non-native speaking teachers of English do not seem to make use of the criterion of intelligibility, and suggest that more weight ought to be given to this criterion.

Most of the studies seem to imply that there is a positive correlation between different categories of errors and different degrees of intelligibility. Thus, Johansson (1973) proposes that tests in which native speakers are required to correct utterances containing different types of errors should make



it possible to calculate the degree of comprehensibility associated with different types. Hughes and Lascaratou (1982), while admitting that in their own test, there is not a good fit between the linguistic classification of the errors used and their apparent effect on communication as revealed in the judges evaluation and comments, seem to be confident that further research will achieve the aim of establishing 'linguistic error categories which provide more accurate predictions of the effect of particular errors on communication' (p. 180). At this point in time, then, advocating that there is a close correlation between the linguistic status of errors and their effect on intelligibility would seem to involve an act of faith rather than logical argument, let alone empirical evidence.

As regards the issue of the judges themselves, there are a number of practical problems as well. One is that we cannot always find them when we need them, or not as many as we would wish to get. Another and more important difficulty is the inconsistency that may exist between them depending on their linguistic or professional background. In the study mentioned by Hughes and Lascaratou (1982), a comparison was made of the evaluations of a variety of errors provided by three different groups of judges: Greek teachers of English, native-speaker teachers of English, and native speakers of English who were not teachers. It was found that the non-teachers made extensive reference to the criterion of intelligibility in explaining their judgements, whereas the Greek teachers did not; the native English-speaking teachers also used this criterion, but to a much lesser extent than the non-teachers.

Clearly, on the basis of these results, it can be argued that errors are perceived differently by judges who are non-native speaker teachers and those

who are native speaker teachers. However, it would seem unwarranted to conclude, or indeed to complain as Hughes and Lascaratou did, that non-native teachers simply do not make use of the criterion of intelligibility.

As Davies (1985: 66) suggests, the real issue may reside in the fact that the degree of intelligibility of an utterance may well be different for different judges because they do not approach it from the same viewpoint. Presumably, the reason why intelligibility was not referred to by the Greek teachers in their assessments is simply that they did not experience any difficulty in understanding any of the erroneous utterances. Similarly, the experienced native-speaker teacher will be familiar with his students' common errors and might therefore deduce the meaning of their utterances in spite of the errors. The wisdom of using a native-speaker non-teacher is somewhat questionable however. An utterance may be easily understood by a non-native speaker teacher who knows both the learner's  $L_1$  and his general learning strategies; but the utterance may completely elude someone who is not used to communicating with non-native speakers of his language. Thus, although, ideally the learner should aim at making himself intelligible to ordinary native speakers, we would suggest that it is not realistic to expect the average non-native speaker teacher of an  $L_2$  to recognise the problems an ordinary native speaker would have in understanding utterances which, to the non-native teacher, may be erroneous but nevertheless intelligible.

This having been said, it would be wrong to argue that intelligibility should not be used in the evaluation of learner's errors. It is certainly worthwhile to try and account for the learner's efforts to communicate, but this should be done with caution. In particular, the main problem is that an erroneous utterance may seem unintelligible in one given context to one given person

and not to other people in other contexts. In view of this difficulty in establishing error gravity we did not rely on interrater evaluation; but, instead, we regarded frequency of error as the main heuristic procedure, with a view to assess the learner's strategic use. In sections 7.7-7.10 a detailed analysis of the learner language behaviour, as an Interlanguage, is presented. We propose that ideally the learner's performance should not be referred to as erroneous since according to Selinker (1972) his attempted production of an L<sub>2</sub> should be more appropriately referred to as a separate linguistic system in its own right. However, in view of its pedagogical significance, learner language can be operationally dealt with in terms of errors and non-errors, following a suggestion made by Zydattiss (1974, 1976).

We also echoed the now famous proposal by Schachter (1974) that there may be 'an error in error analysis'. This statement was originally made in the context of Schachter's (1974) investigation of the avoidance strategy in SLA. In this investigation, Schachter argues in favour of what she calls the 'a priori' approach to contrastive analysis (CA) to predict learning problems that the Chinese and Japanese learners have with relative clauses, and therefore avoid them. This fact, she argues, cannot be accounted for adequately within an error analysis, i.e. since it favours the 'a posteriori' (or explanatory) approach to CA. Essentially, we would agree with Schachter on the need to tackle the avoidance phenomenon among other important learning strategies such as overgeneralization, transfer, risk-taking, etc. We are less convinced, however, by the suggestion that the learning difficulties faced by the Chinese and Japanese learners could not be found among other L<sub>2</sub> learners from totally different backgrounds (Pavesi, 1986). Our own query, however, addressed itself to a different issue, viz. the fact that most standard error analyses are still based on the assumption that, if one had a more direct access to the psycholinguistic

processes, one could allocate errors' causes to one of the two internal causes (i.e. transfer or generalisation). In other words, they suggest that cognitive processes are amenable to quantification. We argued that this would trivialize the whole issue of the complex role of cognitive process in language learning, since in many cases, both internal causes are responsible for the same error type. To this extent, then, there may still be an error in the way error analysis is construed or conducted, and thus, we advocated a qualitative rather than quantitative analysis of data.

Our analysis of the learner's language was conducted by scrutinising the different errors and attempting to find an explanation for them. For convenience, we looked at those errors related to the noun phrase, then the verb phrase, and finally the prepositional phrase. All errors and non-errors as well as the total number of attempts were taken into consideration (i.e. 10, 300 word - sample in total). In the area of the noun phrase, 7 major categories were identified (See Table 7.7) accounting for 32.4% of errors, 54% of non-errors and 51% of all attempted uses in the sample. Basically the difference between levels was found to exist on the quantitative, rather than on the qualitative side. In other words, the learners at different levels of their training in English commit or attempt to produce the same forms, but in varying numbers. In addition, there was no evidence of some groups of learners adopting the avoidance strategy whereby they failed to produce certain forms among these categories (although this was found to occur as we will see later in the verb phrase, for instance). By and large, overgeneralisation, transfer and 'overcompensation' (to use Zydattiss's (1976) term for risk-taking) appeared to characterise the learner's IL behaviour in the NP environment.

In the verb phrase environment, five categories or errors were identified.

(see Table 7.8); and these represented 40.5% of all errors, 18% of non-errors and 21% of all the attempted output in the sample. Here, however, the difference between levels was both quantitative and qualitative (i.e. varying numbers and strategies per level), suggesting that L<sub>2</sub> learners do not necessarily utilize the same strategies to build up the mastery of the rules involving different subsystems of the target language. In other words, language learning may proceed through somewhat different paths in different linguistic environments. Generally speaking, the main learning problem resides in this VP environment, in the correct marking of the irregular past, and the progressive aspect. Learners, at different levels, were adopting roughly similar strategies to deal with these difficulties; but used them in varying proportions. The most frequently used of these strategies was 'overgeneralization'. Unlike with the NP related performance, little evidence of transfer was found. In addition, the strategy of avoidance was regularly resorted to by the university students (i.e. level 2) this leading to their making errors less frequently than the other two groups, in this linguistic environment.

As regards the prepositional phrase (PP), four meaning types were identified involving eight different prepositions (see section 7.10 and Table 7.15). These eight prepositions accounted for 27.1% of all errors, 28% of non-errors and 28% of all the attempted uses. Three main factors appeared to affect the learner's performance in the PP environment: 1) the difficulty in differentiating predicates with positional meaning (i.e. stative verbs) from those with destinational meaning (i.e. dynamic or action verbs); 2) interference from their first language (i.e. French is regarded as the main source of transfer); and 3) overgeneralization of the function of certain prepositions to inappropriate contexts. No instance of the avoidance strategy was found in this linguistic environment.

### 7.12. Summary

In this chapter, we have presented work aimed at discovering a reliable and easily applicable index for assessing  $L_2$  global proficiency. Using writing data, we were able to calculate various measures which when subjected to statistical analysis, indicated that an information processing approach to such data is capable of yielding an index of development of the type described here. To this effect, we proposed the notion of communicational capability index (CCI) as a technique for establishing the degree of complexity and correctness manifested by the analysis of the learner's output. We feel that this technique has gone a long way towards resolving some of the issues that workers in the field have often expressed concern about. Some of these questions are:

1. Would the measures we had found earlier to be successful for discriminating among writing abilities of ESL learners, be equally applicable to oral data?
2. Do the performance variables or measures we have identified increase over time when the subjects are receiving ESL instruction?
3. Is it possible that the proposed index is not immune to influence from the  $L_1$ ? In other words, it might be the case that a person's tendency to write long information blocks in his  $L_1$  can be positively transferred to an  $L_2$ .
4. Can learners be taught to 'beat the system' and write more sophisticated texts than their linguistic abilities permit?

As regard the first question, we think that the combination of average clause embedding weight (ACEW) and the proportion of error-free information blocks (both on which the CCI is based) would equally be applicable to both written and oral data. By contrast, the T-unit concept, as originally devised by Hunt (1965), aims at dealing with written data only. Secondly, our results indicate that the ACEW was the best predictor for the ability to write a complex composition and did increase over time (or level of proficiency). The CCI is

simply the ACEW plus the 3-star credit system to account for increasing accuracy (i.e. EFIB). Thirdly, since the ACEW seems to discriminate well between complex and simple texts written by native speakers (Cook 1979) and L<sub>2</sub> learners, we can argue that it is reasonably immune to L<sub>1</sub> influence. Further investigation, however, is necessary to assess the effect of strategic use on the ability to write more sophisticated texts as a result of instruction. On this fourth issue, then, our results give a mixed finding. However, in an educational setting, it is highly desirable to encourage learners to take risks.

Another important aspect of our analysis of written data addressed itself to the status of learners' errors as part of their learning process. Both quantitative and qualitative methods were used to deal with the issue of how cognitive processes appear to underlie learner language behaviour. Such processes are central to language learning; and their interaction with certain external factors (especially teaching techniques and socio-psychological factors) may lead to the learners committing errors or avoid committing them (i.e. non-errors and avoidance). We found that the main difference across levels was quantitative rather than qualitative. However, the university students, in one particular linguistic environment (i.e. verb phrase) exhibited somewhat qualitatively different performance, resorting to the avoidance of certain structures. This behaviour seems to be at variance with Schachter's (1974) proposal to use a priori contrastive analysis to predict learning problems as evidenced by avoidance of structure. Could contrastive analysis predict the forms that have been avoided by these students? After all, these university students share a common linguistic background with the other Zairean students under investigation. This finding seems to suggest, instead, that as learners progress in their mastery of an L<sub>2</sub>, they may adopt varying strategies at a particular stage and within a specific environment. It also accounts for the fact that

'learning plateaux' are an integral part of the learning process.

Finally, most learners appear to have gone beyond the beginner's stage in their knowledge of English syntax. Indeed, many can be said to be upper-intermediate. We feel that most of the difficulties encountered by these learners could be ironed out provided adequate teaching materials are implemented. More specifically, these materials should be designed using a carefully worked out pedagogic grammar based on relation in discourse.



## CHAPTER 8

### THE RELATIONSHIP BETWEEN THE LEARNER FACTORS AND LANGUAGE

#### PROFICIENCY

##### 8.1. Introduction

In this chapter we shall consider some of the reasons why individual learners study English in the Zairean educational system, and the effect that these learner factors may have on L<sub>2</sub> proficiency. In dealing with the question: 'Why do Zairean learners study English?', at least three reasons may be identified. The most obvious of these is because English is one of the subjects on the curriculum, regardless of the historical sequence of events that led to its inclusion (e.g. rather than Dutch, Spanish, Swahili or Lingala). Then there is the largely unspecified need for international or inter-regional communication between other countries for political, cultural or commercial purposes (Mbaya, 1983). However, one wonders whether this need is truly felt by many Zairean pupils given the high school drop-out rate, which means that for most students the chances of engaging in any significant interaction with foreigners are very slim. But educationists, language planners and the public generally may argue that, at the macro level, the nation must avail itself with adequate means for attaining this general aim, if or when the need arises in changing circumstances. The third reason, then, must involve the micro-level of the classroom learning (i.e. for general or specific purposes) and teaching, whereby both the learners and teacher try to make the most of these choices which somehow have been made for them at the national level. It is this third (micro) level of enquiry that is our main concern in this study. We shall first examine to what extent the learners' own reasons for English study compare with those of the educationalists. In other words, what constitutes in psychological terms,

the learner's orientation motive and attitudes towards the language, its speakers and cultural correlates as well as the learning task itself. Then we shall consider whether these reasons can be regarded as 'predictors' for language achievement. Thirdly, we will establish which learner factors are more likely to have an impact on L<sub>2</sub> progress and eventual success.

## **8.2. A brief review of studies on the influence of the learner factors on Language Proficiency.**

Studies in the literature suggest that L<sub>2</sub> learners vary a great deal depending on such factors as motivation, attitude, learning style, aptitude and age, etc. (Gardner and Lambert 1972; Jakobovits 1970; Spolsky 1969). According to Ellis (1986), individual learner factors fall into two main categories: personal factors and general ones. Personal factors are highly idiosyncratic features of each individual's approach to learning a L<sub>2</sub>. These are exemplified in reports of diary studies (Schumann 1977; Bailey 1980, 1983), in which individuals keep a record of their learning experiences, techniques, and views on teaching methods / materials. General factors on the other hand include variables such as motivation, attitude, age, aptitude, cognitive style and personality. For the purposes of this study we will focus our attention only on studies concerned with motivation and attitude as well as length of L<sub>2</sub> exposure, since these were the factors included in our survey.

The study of motivation and attitude in L<sub>2</sub> learning has attracted many researchers over the last three decades or so. However, one of the major problems remains how one defines motivation and attitude, let alone to pinpoint their role in L<sub>2</sub> achievement. The initial study in this area was conducted by Gardner and Lambert (1959), who administered a number of attitudinal and motivational variables to a group of grade 11 students of French

in Montreal. A factor analysis of the relationship among measures of language aptitude, attitudinal and motivational characteristics, and teacher ratings of students' proficiency in French resulted in the extraction of four factors. Two of these factors were associated with the French proficiency. One of the two factors was defined by the indices of language aptitude, thus supporting the conclusion that achievement in French was related to individual differences in language aptitude. The second factor revealed notable loadings from measures of attitudes towards French Canadians, motivational intensity to learn French, and integrative orientation towards language study. This suggests that achievement in a  $L_2$  was associated with a willingness or desire to learn the language of a valued  $L_2$  community for the purposes of improving communication with them. These two factors (i.e. aptitude and attitudinal characteristics) were orthogonal to each other which means that achievement in  $L_2$  is related to these two independent components.

Since Gardner and Lambert's (1959) pioneering study, several researchers have carried out similar investigations (but to my knowledge, no one has come up so far with such a clear cut pattern) both in Canada (Gardner 1960; Feenstra and Gardner 1968; Gardner and Smythe 1975; Smythe, Stennet and Feenstra 1972; Gardner, Smythe and Clement 1979) and in the United States (Gardner and Lambert 1972; Oller and Perkins 1980; Oller 1983). In addition, other studies have focused on other languages as well as attitudes to particular teaching techniques or programmes.

Anisfeld and Lambert (1961) focused on the relationship of attitudes to learning Hebrew among students in Montreal. Meanwhile Gardner and Santos (1970) studied the relationship between attitudes and achievement in English among senior high school students in Manila, Republic of the Philippines. These

studies were in agreement in demonstrating a relationship between attitudinal characteristics and second language achievement.

A study by Lambert, Gardner, Barik and Tunstall (1963) investigated students attending a six-week French summer school in Montreal. The students were attending an intensive language training program which involved their active participation in the second language for the greater part of the day. The results of this study were mixed, i.e. agreeing only partially with those obtained in previous studies. For instance, for elementary level students, achievement in French was associated with favourable attitudes towards the other cultural group and an integrative orientation toward language study. However, similar relationships were not obtained for the advanced students. Moreover the intensive language programme had the effect of modifying some attitudes. Measures of authoritarianism and anomie increased significantly for elementary level students, whereas for advanced students there was a significant increase in anomie. Additionally, a study by Gardner, Smythe and Brunet (1977) involved high school students registered in a five week intensive French programme. They found that over the course of the programme, students tended to become more ethnocentric, less interested in foreign languages and less integrative in their reasons for language study. Thus, according to these results, an intensive language programme can make individuals less tolerant than before towards the  $L_2$  and its speakers' community. Other results show that students become more comfortable with the language (i.e. less anxious in the classroom situation) as a result of the intensive programme.

On the other hand, Gardner, Smythe and Clement (1979) set out to determine the relation of attitudinal / motivational attributes to second language achievement, and the effects of an extensive language programme on

attitudinal or motivational characteristics and L<sub>2</sub> proficiency. Two different samples were involved, one Canadian, and the other American. A factor analysis of the 24 variables demonstrated an association between an attitudinal / motivational factor (or integrative motive as they call it) and a French achievement factor. The results also showed satisfaction with the programme for the Canadians but not the Americans. A third factor was referred to as anxiety. Both samples evidenced decreases in anxiety and attitudes towards bilingualism and increases in French proficiency as a result of the programme. These results indicate the role played by attitudes in L<sub>2</sub> achievement, in particular, the influence that the socio-cultural background of the student can have on the eventual success in L<sub>2</sub> learning. Others have emphasized the importance of length of residence in the target language community (Eskstrand 1975; Hatch 1983) or length of L<sub>2</sub> exposure and learner's contribution (Seliger 1983; Upshur and Homburg 1983; Ramirez 1984; Politzer and McGroaty 1983) as the best predictors for L<sub>2</sub> achievement. Others still, such as Oller, Perkins and Murakami (1980), have questioned the validity of the attitude surveys on grounds that the self-report questionnaire may constitute an unintentional language and intelligence test. However, if sufficient care is taken both in the elaboration of the questionnaire and interpretation of the results, we believe that the study of the learner factors and their relationship to L<sub>2</sub> achievement are of great importance for both theoretical and educational ends.

### **8.3. Experimental Design**

The questionnaire used for investigating the learners' attitudinal characteristics consisted of two main sections. Section One was aimed at eliciting demographic information i.e. in relation to the subject's identity, sex, educational and sociolinguistic backgrounds, as well as a self-assessment of proficiency in both French and English. Owing to limitation of time, this type of

information is not reported in the present study.

Section Two of the questionnaire focused more specifically on the attitudinal attributes of the respondents and consists of three parts (see Appendix for details on the items and instructions given). The three parts are entitled as follows:

1. Zairean learners' attitude towards English (Attitude 20 items)
2. Zairean learners' orientation index (Motivation: 10 items)
3. Zairean students' desire to learn English (Desire to learn: 10 items)

The questionnaire for Section Two was constructed in such a way as to reflect a wide range of responses and implied 15 variables. These were correlated with three language based variables totaling 18 variables as follows:

1. **Ethnocentrism** – This reflects the belief that the Zairean culture(s) could be under threat if European ways became the norm (e.g. through marriage between Zaireans and foreigners: cfr items 14, 18 on attitude). A low score means a higher degree of ethnocentrism.
2. **Anomie** – The items implying anomie (i.e. 5 and 7 on attitude) assess the subjects' tolerance or willingness to adopt Western habits or patterns of behaviour. The higher the score the greater the feelings of anomie.
3. **Perceived Utility of English** – These statements demonstrated how useful the students think the English course is, e.g. for studying Maths and Science subjects; and whether the knowledge of English can improve their self-images as 'educated people' (items 1, 4 and 19 on attitude).
4. **Attitudes towards the English Language and its Speakers' Culture** – These scales contain positively worded statements suggesting that the British people should be proud of their language and culture; the scales also urge young Zaireans to make a positive effort to meet English speaking persons (items 3, 6, and 12 on attitude).
5. **Attitudes towards Native vs Non-Native teachers of English** – This scale assesses the students' preference or attitude towards native speaker teachers as well as the existence of anti-foreigner feelings (item 10 on attitude)

6. **Contribution of the English speaking people to Zaire's development** – These scales consist of positively worded statements to assess whether the respondents believe the English speaking expatriates in Kinshasa contribute significantly to the country's development (items 11, 12 and 13 on attitude).
7. **Interest in Foreign Languages (English or French)** – These statements assess the wishes of the respondents to obtain more input or exposure to English and/ or French (items 15, 16 and 17 on attitude).
8. **Desire to Socialize with the Expatriates' Community** – These scales consist of positive statements to indicate the students' preference of the expatriates they would choose to socialize with. The choices include Americans, Belgians, British, French, Japanese, Lebanese and others ( items 6, 8, 9 and 20 on attitude).
9. **Ratings of Instrumental Orientation** – These five scales consist of statements involving utilitarian reasons for studying English, e.g. getting a well-paid job or higher marks at school / college (items 1, 3, 5, 7 and 9 on motivation).
10. **Ratings of Sentimental Orientation** – These five scales emphasize the students' need to meet or correspond with English speaking friends for the purpose of promoting cultural understanding (items 2, 4, 6, 8 and 10 on motivation).
11. **Motivational Intensity** – This is a 0 – 10 point scale indicating the students' rating of the English course, compared to 10 other subjects or courses in the curriculum (item 1 on desire to learn).
12. **Desire to Develop the Communicative Use of English** – The students are asked to state whether greater linguistic knowledge would increase their willingness to read novels or short stories in English (item 4 on desire to learn).
13. **Attitude towards Instruction I** – These scales assess whether the students are satisfied with the amount of contact hours devoted to English in the timetable. Should the number of hours be increased or decreased? (items 6 and 7 on desire to learn).
14. **Attitude towards Instruction II** – These scales consist of self-reports of the students' endeavour to participate in classroom learning (items 3 and 8 on desire to learn).
15. **Perserverance** – These scales assess how committed the students are to do their homework or to continue to learn when they leave school / college; e.g. by joining an English club (items 2, 5 and 10 on desire to learn).
16. **Modified Cloze Test** – This is a researcher-made reading passage to assess the use of noun phrase reference; i.e. all articles are omitted from the passage and the learners are asked to supply them where they think it is appropriate to do so (74 items).

17. **Sentence Interpretation** - This task consists of 36 sentences in relation to the identification of the correct subject noun phrase of an infinitival complement clause.
18. **Overall L<sub>2</sub> Communicational Capability Index** - This measure involves the assessment of a (Picture description) written composition in terms of: 1) the average clause embedding weight (Cook 1979), and 2) the percentage of error-free information blocks (see Chapter 6).

#### 8.4. Subjects

Subjects for this survey are seventy-one students learning English at high school, university and teachers' college levels. Although the questionnaire was distributed to all of the 163 students who took part in the main study, more than half did not return them to their teachers as requested. The actual number of subjects was: 35 for high school; 11 for university and 25 for the teachers' college. Their age varied between 17 and 22 years. Sex and socio-economic background were not controlled for.

#### 8.5. Materials and Procedure

A total of 40 items were included in the questionnaire (see Appendix). As already indicated, the items consisted of Likert-scales; i.e. showing degrees of agreement with (4 and 6) alternative statements. These choices may indicate either: strong support (A), moderate support (B), little support (C), little disagreement (D), moderate disagreement (E), or strong disagreement (F). To save valuable classroom time, the questionnaire was handed out to the students for completion at home.

#### 8.6. Results

The fifteen scales implied in the questionnaire are impressionistically grouped into four major variables. These are; 1) attitude, 2) sentimental motivation, 3) instrumental motivation, and 4) desire to learn English.



Additionally there are four language based variables: 5) modified Cloze test, 6) sentence interpretation, 7) average clause embedding weight, and 8) error-free information blocks. Three types of results are given: a) Product - moment correlations, b) Factor analysis, and c) Anova and Scheffe tests. Types (a) and (b) are given together for convenience.

### 8.6.1. Correlations and Factor Analysis

First of all, we examined the relationship between four learner variables and the more grammatically -based  $L_2$  tasks (i.e. modified cloze test and sentence interpretation). Table 8.1 gives the correlation matrix for six variables. These results show that there is a strong positive relationship between measures of attitude and both sentimental and instrumental orientation motives ( $p < .01$ ), as well as a positive but weaker relationship between these three and desire to learn English ( $p < .05$ ). However, no relationship is found between the learner variables (i.e. variables 1 to 4) and measures of language proficiency ( i.e. 5 and 6).

Table 8.1

**Pearson Correlation Matrix for Attitude/Motivational Survey  
and Two Language Tasks**

(N = 71 subjects)

VARIABLES	1	2	3	4	5	6
	ATT	SENT.M	INST.M	DESLN	CLOZE	S. IN.
1. Attitude	--					
2. Sentimental Mot.	.50**	--				
3. Instr. Mot.	.53**	.61**	--			
4. Desire to Learn E..	.35*	.26*	.12	--		
5. Mod. Cloze	.14	.12	-.04	.17	--	
6. Sen. Inter.	-.14	-.14	-.12	.02	-.01	--

\*\* p < .01  
\* p < .05

In order to establish the underlying traits among the above variables, the correlation matrix was factor analysed by means of a Principle Factor solution (PA2). Four factors were extracted; but application of the Scree Test (Cattell 1966) showed that only one factor had eigenvalues greater than 1.00, whereas the other three factors were very close to 1.00 ( i.e. 0.98).

Table 8.2

## Unrotated Principal Factor Solution [PA2] With Iterations

VARIABLES	FACTOR (1)	FACTOR (2)	FACTOR (3)	FACTOR (4)
1. Instr. Motiv.	.84	-.40	.19	.24
2. Sentimental Mot.	.67	-.05	.11	.11
3. Attitude	.67	-.01	.01	-.01
4. Cloze Test	.23	.73	-.40	.49
5. Sent. Inter.	-.21	.39	.87	.22
6. Desire to learn English	.54	.53	.07	-.64

In view of the appreciable loadings revealed in the unrotated matrix (Table 8.2), the four factors were rotated using the varimax procedure (Kaiser 1958).

The rotated factor matrix is given in Table 8.3. Factor 1 obtains positive and substantial loadings from instrumental and sentimental motivation and attitude. This pattern suggests that we are dealing with an attitudinal / motivational factor. According to these results, students who advocate utilitarian reasons for studying English (variable 1) also evidence great interest in establishing friendly relationships with English speaking people (variable 2). In addition, students with strong instrumental orientation also have highly favourable attitudes towards English speaking people and their culture (variable 3). However, unlike in the Gardner and Lambert (1959) study, none of these attitudinal characteristics is related to L<sub>2</sub> proficiency measure. Instead, these findings are similar to those obtained by Oller, Perkins and Murakami (1980) who found no clear pattern between measures of attitudes towards Americans (their subjects were 182 foreigners in the USA) and L<sub>2</sub> proficiency.

Table 8.3

## Factor Matrix (VARIMAX rotation with Iterations)

Four subtests of the Attitude / Motivation Survey, a modified Cloze test, and sentence interpretation task. n = 71.

VARIABLES	FACTOR (1)	FACTOR (2)	FACTOR (3)	FACTOR (4)
1. Instr. Motiv.	.97	-.07	-.10	-.05
2. Sentimental Mot.	.67	.14	.08	.01
3. Attitude	.61	.24	.09	-.09
4. Cloze Test	.06	.08	.99	-.01
5. Sent. Inter.	-.07	.01	-.01	.99
6. Desire to learn English	.19	.98	.08	.02

Factor 2 obtains substantial loadings from desire to learn English (variable 4) and positive but small loadings from attitude (variable 3). These results suggest that students who evidence a strong desire to take an active part in their formal instruction in English also tend to show favourable attitudes towards the English speaking people and their culture. This factor represents what could be called motivational intensity towards instruction. In addition, there is no relationship between this factor and L<sub>2</sub> grammar based tests. Factor 3 demonstrates that the modified cloze test represents a separate factor, since practically all the variance for this variable is accounted for (i.e.  $r^2 = 98.8\%$ ). Finally, factor 4 represents the other grammar based test, i.e. sentence interpretation. As with factor 3, almost the total amount of variance is explained by this variable alone (i.e.  $r^2 = 99.4$ ). These results so far suggest that none of the learner variables constitutes a predictor for L<sub>2</sub> proficiency.

Since the two grammatically oriented tasks (i.e. modified cloze and sentence interpretation tests) did not correlate with the attitudinal variables, we decided to include the more communication based measures representing a global index of  $L_2$  proficiency viz. the average clause embedding weight (ACEW) and the proportion of error-free information blocks (EFIB). Unfortunately, this led to the reduction of the number of students involved in the correlational analysis from 71 to 33, since some of those who did return the questionnaire had not produced compositions which were long enough to be included in the analysis of the latter.

The same factor analysis procedure as used above was followed for this sample. First, the correlation matrix (Table 8.4) shows clearly that attitude, sentimental and instrumental motives and desire to learn English are positively related (values range between  $r = .76$  and  $.30$ ). However, there is little in common between these learner variables and  $L_2$  proficiency. On the other hand, there appears to be a positive and systematic relationship among most of the language-based variables, i.e. error-free information blocks correlate with modified cloze test and clause embedding weight.

Table 8.4

**Pearson Correlation Matrix for Attitude / Motivation  
Survey and Four Language Tasks**

VARIABLES	ATTD (1)	SENT M (2)	INST M (3)	DESLN (4)	CLOZE (5)	S INT (6)	CL EMB (7)	ERR FREE (8)
1. Attitude	---							
2. Sentimental Mot.	.59**	---						
3. Instrument Mot.	.65**	.76**	---					
4. Desire to Learn	.30*	.41**	.12	---				
5. Mod. Cloze	-.04	.09	.02	.11	---			
6. Sen. Interpre.	.05	-.10	.10	-.13	.10	---		
7. Clause Embed.	-.21	-.28	-.13	-.07	.10	.06	---	
8. % Error-Free Info Blocks	.07	.06	-.09	.30*	.55**	-.04	.49**	---
** p < .01								
* p < .05								

Secondly, six factors were extracted using the Principal Factor solution (PA2) given in Table 8.5. This matrix was then rotated by the Varimax procedure (Table 8.6). The results demonstrate the existence of an underlying attitudinal factor (1) with heavy loadings on the instrumental and sentimental motive and moderate ones on attitude. This means that students who believe strongly in utilitarian reasons for learning English (variable 1) also show a high integrative or sentimental orientation (variable 2). Additionally they have favourable but moderate attitudes towards the English speaking people and culture. Factor 2 is identified by two variables (3 and 4), indicating that students who have the

Table 8.5

## Unrotated Principal Factor Solution [PA2] With Iterations

Four Subtests of the Attitude / Motivation Survey and Four L<sub>2</sub> Performance Tasks (n = 33)

VARIABLES	FACTOR (1)	FACTOR (2)	FACTOR (3)	FACTOR (4)	FACTOR (5)	FACTOR (6)
1. Instr. Motiv.	.86	-.06	.39	.09	-.19	.26
2. Attitude	.84	.01	.14	.14	.09	-.45
3. Sentimental Mot.	.79	.05	-.04	-.01	-.12	.13
4. % Error Free Blocks	.01	.75	-.04	.05	-.03	-.11
5. Cloze Test	.05	.75	.09	-.60	-.18	-.01
6. Av. Clause Embedding	-.37	.62	.33	.60	-.03	.06
7. Sent. Inter.	-.02	.02	.57	-.22	.57	.07
8. Desire to learn English	.48	.39	-.61	.11	.40	.16

Table 8.6

## Factor Matrix (Varimax rotation with iterations)

For the Four Attitude / Motivation Survey and Four L<sub>2</sub> Performance Tasks (n = 33)

VARIABLES	FACTOR (1)	FACTOR (2)	FACTOR (3)	FACTOR (4)	FACTOR (5)	FACTOR (6)
1. Instr. Motiv.	.99	-.04	-.01	-.04	.10	.12
2. Sentimental Mot.	.73	.08	-.15	.25	-.10	.18
3. Cloze Test	.06	.98	.02	.01	.08	-.07
4. Error Free Blocks	-.04	.56	.44	.24	-.07	.12
5. Clause Embed.	-.13	.10	.98	-.05	.04	-.07
6. Desire to learn English	.17	.11	-.01	.94	-.09	.09
7. Sent. Inter.	.01	.04	.02	-.07	.84	.02
8. Attitude	.57	-.02	-.06	.15	.05	.78

greater ability to use the articles correctly (i.e. aspects of syntax and semantics) are also more likely to produce error-free sentences in a written composition. This factor could be referred to as productive grammatical accuracy. None of the learner variables (factor 1) could be regarded as a predictor for this  $L_2$  performance area.

Factor 3 is defined by two measures of overall  $L_2$  proficiency. The results indicate that students who are able to produce more elaborate compositions (variable 5) also evidence greater grammatical accuracy (variable 4), albeit moderately so. This factor seems to represent the learners' overall communicational capability in English.

Factor 4 reflects the motivational intensity towards instruction. This factor obtains heavy loadings from desire to learn (variable 6) and smaller ones on sentimental (but not instrumental) orientation (variable 2), as well as error-free information blocks (variable 4). This factor implies, surprisingly, that the Zairean students who have strong motivational intensity towards English language study do so, not for utilitarian reasons, but for sentimental ones. And yet at the University of Kinshasa, for instance, we have always taken it for granted that at least our Science students will strive to learn English, given the utilitarian ingredient that this may involve for their studies. We will return to this later on (see section 8.7).

Factor 5 consists of the sentence interpretation task on its own. Performance on this task seems totally unrelated to other aspects of the students'  $L_2$  performance. This factor could be called 'receptive grammatical accuracy' and involves quite separate skills on the students' behalf. Finally, factor 6 receives appreciable loadings from attitude (variable 8), implying the presence of an orthogonal attitude factor different from the one we described



above (factor 1). Further investigation especially with larger samples might take issue with this theoretically unusual trend. However, this may well be a purely methodological matter. According to Carroll (1983: 105) overfactoring is indicated by a solution in which one or more factors contain apparently significant loadings on only one variable, whose composition is already well accounted for by solutions with a smaller number of factors. According to Carroll, such apparently significant loadings are likely to represent specific or error variance (i.e. rather than common variance). Thus, we regard factor 6 as being an artifact resulting from either the content and sample sizes or the computational procedures involved. Indeed runs were made assuming either three, four or five factors; but the computations were terminated early in the iteration process, i.e. before the varimax rotation could converge or stabilize.

#### **8.6.2. Analysis of Variance and Scheffe Tests**

We now turn to the question whether the learner variables or characteristics themselves remained constant or varied in relation to the length of exposure (i.e. English training levels). To do this we must compare the means for:

1. Instrumental Orientation Motive
2. Sentimental Orientation Motive
3. Attitude towards English speaking people and culture
4. Desire to learn English
5. The four language-based tasks (as in Chapters 5, 6 and 7) by level of training.

The ANOVA results for the instrumental motivation are given in Table 8.7. They indicate that the pupils evidence a higher amount (65.86%) than both the teacher trainees (53.8%) and the university students (48.6%). These results are highly significant too ( $F = 6.75$ ;  $p < .002$ ). Appropriate Scheffe tests (Table 8.7)

demonstrate that these differences are highly significant ( $p < .005$ ). This means that as the students progress in their training their outlook or reasons for learning English changes considerably. In their beginning years, the students perceive language learning in utilitarian terms (e.g. getting higher marks or a well paid job). However, after a certain time, they may realize that knowing the language is an advantage but not the only factor involved in reaching these goals.

Table 8.7

## ANOVA – Instrumental Motivation by level of Training in English

N =71

LEVEL	N	MEAN	SD
1. Pupils	35	65.86	13.03
2. Undergrads	11	48.64	13.62
3. Teacher tr.	25	53.80	20.42
F = 6.75		P < .002	
SCHEFFE TESTS	L2 Undergrads	L3 Teacher Tr.	L1 Pupils
X =	535.04	1345.0	2305.1
L2 undergrads	-----	809.96***	1770.06***
L3 teacher tr.		-----	960.1***
*** P < .005 t crit = 311.76			

As regards the sentimental orientation (Table 8.8) the means for the three levels are quite high (around 75%); but the ANOVA results do not reach significance ( $F = 0.13$ ; n.s.). These results mean that these Zairean students' orientation motive is largely sentimental and remains unchanged during the best part of their formal instruction or beyond. This finding is in disagreement

with that of Lambert, Gardner, Barik and Tunstall (1963) who found that (after a six-week intensive course in a French summer school in Montreal) some of the learners attitudes changed significantly. Measures of authoritarianism and anomie increased for the elementary level; whereas for the advanced students there was a significant increase in anomie. In other words, intensive courses can make individuals become less tolerant of the L<sub>2</sub> cultural correlates, whereas extensive language courses do not.

**Table 8.8**

**ANOVA – Sentimental Motivation by Level of Training in English**

N =71

LEVEL	N	MEAN	SD
1. Pupils	35	75.91	8.21
2. Undergrads	11	76.36	12.67
3. Teacher tr.	25	74.80	11.23
F = 0.134		P = n.s.	

Moreover, the results for attitude towards English language and culture (Table 8.9) indicate that these students' attitudes are favourable, but only moderately so (i.e. 60 % on average). Additionally no change is recorded over time across the three different levels ( F = 0.497; n.s.). Gardner, Smythe and Clement (1977) found that the Canadian students' attitudes were positively affected by their extensive language programme, and they went even further to suggest that this may lead to an increase in L<sub>2</sub> proficiency, a fact which is not supported by our data. On the other hand, measures of desire to learn English appear to be the most significantly affected by the students' level of training (Table 8.10). We find that the means for the teacher trainees is the highest (i.e. 81.02 %), then

the university students (i.e. 77.5%), and finally the pupils come in third position (i.e. 72.97%). The ANOVA results demonstrate that these means differences are significant ( $F = 3.53$ ;  $p < .005$ ); and so are all possible Scheffe tests for post hoc comparisons. This indicates that the students' attitude and personal involvement in their instruction increased considerably as they received more and more training or exposure to English. However, this may not be interpreted as an indication of a cause to effect relationship between length of exposure and desire to learn. Presumably those who have enjoyed earlier success will go for yet more success; thus they evidence an increasing desire to learn.

**Table 8.9**

**ANOVA – Attitudes towards English Language/Culture by training in English**

N =71

LEVEL	N	MEAN	SD
1. Pupils	35	63.06	13.68
2. Undergrads	11	58.55	10.49
3. Teacher tr.	25	60.88	15.18
F = 0.497		P = n.s.	

Table 8.10

## ANOVA - Desire to Learn English by Level of Training in English

N =71

LEVEL	N	MEAN	SD
1. Pupils	35	72.97	13.76
2. Undergrads	11	77.55	8.40
3. Teacher tr.	25	81.02	9.33
F = 3.53		P < .005	
SCHEFFE TESTS	L2 Undergrads	L3 Teacher Tr.	L1 Pupils
mean X =	835.05	2025.5	2553.95
L2 undergrads	----	1172.45***	1700.9***
L3 teacher tr.		----	528.45***
*** P < .005 t crit = 266.38			

Finally, as far as the language based tasks / variables are concerned, the effect of length of exposure was mixed. No significant differences are found across levels regarding the more grammar-oriented tasks; viz. the modified cloze test and sentence interpretation. On the other hand, the three levels of training differ significantly (  $p < .005$ ) in relation to their global index of  $L_2$  communicational ability, i.e. in terms of the degree of clause embedding and proportion of error-free information blocks.

Table 8.11

**ANOVA - Performance on the Modified Cloze Test  
by Level of Training in English**

N = 71

LEVEL	N	MEAN	SD
1. Pupils	35	62.06	11.64
2. Undergrads	11	62.45	14.65
3. Teacher tr.	25	62.80	12.57
F = 0.026		P = n.s.	

Table 8.12

**ANOVA - Performance in the Sentence Interpretation Task by Level**

N = 71

LEVEL	N	MEAN	SD
1. Pupils	35	72.69	20.7
2. Undergrads	11	71.09	14.29
3. Teacher tr.	25	69.08	15.77
F = 1.28		P = n.s.	

Table 8.13

ANOVA - Average Clause Embedding Weight by Level of Training  
in English

N =71

LEVEL	N	MEAN	SD
1. Pupils	57	1.47	0.18
2. Undergrads	11	1.58	0.50
3. Teacher tr.	35	1.68	0.24
F = 6.88		P < .002	
SCHEFFE TESTS	L2 Undergrads	L3 Teacher Tr.	L1 Pupils
X =	17.35	58.63	84.02
L2 undergrads X = 17.35	----	41.28**	66.67**
13 teacher tr. X = 58.63		----	25.39**
** P < .01 t crit = 16.51			

Table 8.14

## ANOVA – Proportion of Error-free Information Blocks by Level of Training

N =71

LEVEL	N	MEAN	SD
1. Pupils	57	0.31	0.16
2. Undergrads	11	0.49	0.20
3. Teacher tr.	35	0.47	0.23
F = 8.90		P < .001	

These results (Table 8.13 and 8.14) are indeed in agreement with the Gardner et al. (1977) study mentioned earlier, suggesting that length of exposure (as in extensive programmes) may be a significant factor leading to eventual success in L<sub>2</sub>.

### 8.7. Discussion

We shall discuss, in turn, the results in relation to 1) the role of the attitudinal variables, 2) the nature or structure of the learners' competence, and 3) the effect of L<sub>2</sub> proficiency. First of all, there is no relationship between the learner variables (i.e. attitude, sentimental or instrumental orientation and desire to learn English) and the two grammatically - oriented tasks (Table 8.1). This finding is similar to that obtained by Oller, Perkins and Murakami (1980), who found that none of the attitudinal variables could predict higher scores for grammar or dictation. In their study, some of the variables which were regarded as extraneous to the study (e.g. being in favour or against abortion, or the legalization of marijuana) accounted for as much variance as the attitudinal variables. Thus, Oller et al. (1980) conclude that attitude questionnaires are a



kind of unintentional language test. They also argue that such questionnaires may well be regarded as tapping the learners' intelligence and general ability to give consistent, socially appropriate and self-flattering responses – rather than leading to an assessment of the theoretical construct involved in so-called attitudes. Since no extraneous variables were included in our own questionnaire, we shall not deal specifically with Oller et al. speculation on the use of self-reporting questionnaires. But if indeed such questionnaires are largely a language test, especially apt to elicit socially appropriate responses, then our subjects must have had it so good, since all the questions were asked in French to make them easier to comprehend. Oller et al. argument seems flawed on grounds that if intelligence was the main construct at stake, then both the attitudinal and  $L_2$  proficiency variables might share some variance in common. At the very least, the rotated factor matrix (Table 8.6) might reveal that all the dependent variables contribute substantially to one distinct factor; whereas the independent (or learner) variables constitute another.

Our results, however, show that this is not the case. What we find instead is that one factor, which we have called 'attitudinal' (i.e. involving instrumental and sentimental orientation, and attitude) is orthogonal to another one, referred to as motivational intensity towards instruction (involving desire to learn). Additionally,  $L_2$  proficiency consists of three distinct factors. However, we are not making claims as regards the nature or structure of the learner's competence, i.e. in terms of the hypothesis of a unitary versus divisible competence (Oller 1979; Oller and Hifonotis 1980; Farhady 1983; Palmer 1983; Canale 1983). This issue has generated a great deal of debate among  $L_2$  researchers over the last decade or so; and the emerging consensus indicates that 'there is a general language ability but at the same time language skills have some tendency to be developed and specialized to different degrees, or at

different rates, so that different language skills can be separately recognized and measured' (Carroll 1983: 82). Our results seem to reflect Carroll's multifactor approach to the structure of language ability, rather than the unitary language competence approach.

In Table 8.6, we see that there is a connection between the 'active productive ability' as measured by the modified cloze test (i.e. variable 3) and the ability to produce correct sentences as measured by the proportion of error-free information blocks (i.e. variable 4). At the same time, the two measures representing the overall L<sub>2</sub> communicational ability (variables 4 and 5), contribute substantial loadings towards the same factor, viz factor 3. On the other hand, factor 5 deals only with linguistic competence, especially 'active receptive ability', as measured by sentence interpretation (variable 7). If we further take into account the fact that, according to the ANOVA results, there is no difference between the three levels on the linguistic competence measures, whereas there are significant differences between levels on the more communication-based variables, we can then argue that these results provide corroborative evidence to the distinctiveness between linguistic competence and communicative competence. Our results also lend some support to the notion of the interdependence of linguistic and communicative competences. Within the context of this study, we regard an increase in the ability to produce error-free sentences as the lower bound of communicative competence, i.e. a reduction of the possibility of breakdown in communication due to 'error gravity' (Enkvist 1973). In addition, we assume that the ability to produce elaborate sentences (involving an increasing degree of embedding) indicates a higher level of communicational competence. If this is so, then the interdependence of linguistic competence and communicative competence is clearly shown by these results (Table 8.6). In this connection, factor 2 obtains

heavy loadings from modified cloze test (i.e. focusing on linguistic forms) as well as moderate loadings from error-free information blocks. Meanwhile factor 4 receives heavy loadings from measures of clause embedding weight and moderate ones from error-free blocks.

The above results are in agreement with those of Politzer and McGroaty (1983) who undertook a correlational study of the communicative competence of Spanish speaking pupils in Bilingual education programmes. Politzer and McGroaty suggest that: first, low levels of linguistic competence appear incompatible with high levels of communicative competence; second, high linguistic competence does not guarantee a high degree of communicative competence; and thirdly, different levels of communicative competence are not possible at the same level of linguistic competence. Politzer and McGroaty (1983) and Ramirez (1984) make the further suggestion that there is a minimum low level of linguistic competence as a prerequisite for adequate communicative competence to take place. In addition, they argue that linguistic competence develops as quite distinctive from communicative competence; and that the former presupposes the latter, whereas the converse does not.

We turn now to the third main point of our discussion, viz. the effect of L<sub>2</sub> exposure on attitudes and on the overall L<sub>2</sub> developments. Among the four learner variables, only two were affected by length of exposure. On the one hand, the instrumental orientation (Table 8.7) and desire to learn English (Table 8.10) show significant change over time. Thus, these are the most likely variables to influence the eventual success in learning English for these Zairean students. On the other hand, sentimental orientation and attitudes towards the language and culture remain constant (Table 8.8, 8.9). This of course does not mean that these two learner variables are not important at all. Indeed

motivation in its widest (if complex) sense and interest in the L<sub>2</sub> (due to general attitudes in the community) are very important for language learners (Corder 1967; McDonough 1981). Fortunately enough, both sentimental orientation and attitudes reveal favourable or positive trends. Let us examine in some detail the various changes that are taking place. First of all, we have just observed that both the instrumental and sentimental orientation variables show different patterns. In their extensive studies, Gardner and Lambert (1972) have investigated to what extent achievement in a L<sub>2</sub> is related either to the desire to use the language in the student's own community for instrumental reasons (i.e. for business, promotion or simply to possess a qualification) or to use it for sentimental reasons (i.e. to become an accepted or valued member of the community that speaks the language). They have found that in places like Montreal, where English and French speaking Canadians live side by side, higher achievement in the other language is associated with the sentimental (i.e. integrative) orientation. Obviously, the definitions of these two orientation types may depend on the kind of questions asked of the learners; and indeed Gardner and Lambert's own conceptualization of these constructs have changed somewhat over the years.

Jakobovits (1970:270) provides probably the most widely accepted questionnaire type and defines these two variables as follows. An instrumental orientation is assumed from the respondents' agreement with these statements:

The study of (English) can be important to me because:

1. I need it to finish secondary school (i.e. higher marks)
2. One needs a good knowledge of at least one foreign language to merit social recognition.

3. I think it will some day be useful in getting a good job.
4. I feel that no one is really educated unless he is fluent in (English).

An integrative orientation to language learning is involved if the learner agrees with the following statements:

The study of English can be important to me because:

1. It will enable me to gain good (pen) friends more easily among the (English) speaking people.
2. It will help me understand better the (English) speaking people and way of life.
3. It allows me to meet and converse with more and varied people.
4. It should enable me to think and behave as the (English) speaking people do (adapted from Jakobovits 1970).

It appears that under the sentimental / integrative orientation motive there are two different aspects: one is social (i.e. a general desire for wider social contact), the other psychological (i.e. a stronger belief that learning a L<sub>2</sub> will let the individual acquire the psychological characteristics of the other group). Thus a sentimental / integrative means that the learner wants either to affiliate himself and/or belong in the community. It is not clear which of these two aspects is intended in the Gardner and Lambert (1972) study, in which they find a positive correlation between the integrative orientation and achievement in L<sub>2</sub> learning. In one of the Spolsky's (1969) findings, foreigners learning English for study in the United States appear to be integratively motivated in the psychological sense. It is unlikely the Zaireans learning English would perceive it in this stronger sense of belonging or wanting to assume personality characteristics which might be described as typical of the British, Americans and other English speaking people in Zaire. First there are very few personal

contacts between the students and member of the native speaker communities. Second, the now established 'Authenticity movement' initiated by the Government in the 70's (to foster the revival of the national identity and indigenous cultures) is strong enough to prevent the feelings of 'anomie' from developing significantly. Thus, we feel confident in proposing that it is the social aspect of sentimental orientation which characterises most Zaireans learning English.

According to our results (Table 8.8) the students at all three levels maintain a high and positive sentimental orientation in the latter, social (i.e. affiliative) sense, not the former, psychological one. However, this sentimental orientation does not correlate with the  $L_2$  dependent variables (Table 8.4). It is therefore an open issue whether the sentimental or integrative orientation is indeed effective for many  $L_2$  learners. In Canada, for instance, Gardner and Smythe (1975) found that an integrative motive is dominant among many learners in Toronto; and they still point to both high achievement figures and drop-out rate. Gardner and Lambert (1972: 121-130) and Lukmani (1972) have found instrumental motivation to be more powerful, in Manilla (Philippines) and Bombay (India) respectively. Our results (in Table 8.7) show that the instrumental orientation rate is lower (only 56.1 % on average) than its counterpart (i.e. 75.7%) for all groups. The interesting point about our results however is that the instrumental (but not the sentimental) orientation rate is affected by the level of training in English. This demonstrates that the learners' outlook or reasons for studying the language are changing. The pupils show the highest instrumental orientation percentage (Table 8.7), then the teacher trainees are second, and the university students have the lowest.

It is clear that in the early stages, learners generally believe that knowing

the language will bring about some obvious material gains. This assumption may lead the learners into being willing to learn the language. However, as they proceed in their instruction they may discover that knowing the language is an asset, but it does not constitute the only condition to be fulfilled in order to obtain a good job. Additionally, the learners may become disillusioned about making substantial progress in English if they are not satisfied with the teaching materials or methods being used. This may explain the lower instrumental orientation rate, especially among the university students, who have only about one hour exposure per week (or a maximum of 30 hours per academic year). Although most are still well motivated they gradually become frustrated about their actual  $L_2$  achievement and eventual success. On the other hand, we see that (Table 8.7) the teacher trainees do not study English for purely instrumental reasons. It is likely that they too are primarily concerned about jobs, so that in spite of the greater amount of exposure and more effective teaching techniques, the teacher trainees are just as anxious as any other student about securing a well-paid job. All these results reach very high significance levels (beyond  $p < .01$  for both ANOVA and Scheffe test), which means that they demonstrate deep-rooted perceptions or anxieties among the learners. Therefore, both educationists and teachers (especially ESP teachers at the university where we find the lowest rate of all) ought to take them into account when allocating the number of contact hours or implementing other aspects of the curriculum.

Next we examine the learners' desire to learn English and how it is affected by the length of exposure. The results indicate that the students' contribution or involvement in formal instruction is high, and increases significantly as they proceed with their studies. This at least seems to be an encouraging finding, since the teachers can feel confident that the learners are willing to co-operate

in the learning process. This firm desire to take an active part in their instruction is equally sustained by favourable attitudes (Table 8.9). These attitudes seem to be due to the widespread prestige enjoyed by the French and English Languages (or Western culture) in the Zairean community at large. In addition, the fact that those students who managed to make progress or to succeed in their formal education also show growing interest in English classes brings fresh evidence to Burstall's (1975) comments that 'in the language learning situation nothing succeeds like success'. It is interesting to notice that the learners' desire to learn is, in fact, the only learner variable which (in the present study) serves as a 'predictor' for  $L_2$  dependent variables, viz. the proportion of error-free information blocks and the degree of embedding in subordination (Table 8.4). In pedagogic terms, this suggests that learner-centred methods constitute the most appropriate teaching technique for developing communicative competence and eventual success in the target language. This is also so because if the students are made more responsible, i.e. in charge of the learning task, and if they too enjoy just that, we can be confident that their rate of learning will increase dramatically; with those who have enjoyed earlier successes being also the most likely to achieve more success. In psychological terms, these results mean that past success alone can indeed 'predict'  $L_2$  achievement; but only active learner's involvement (i.e. hard work) can guarantee it.

Finally, let us examine the relationship between the length of exposure and  $L_2$  achievement (Tables 8.11-14). Here we find a mixed pattern, depending on whether the measures are more related to linguistic competence rather than communicative competence. Generally, measures focusing on specific forms do not differentiate between groups; neither do they seem to be affected by the level of training reached (Table 8.11-12). It is somewhat disturbing to notice



that (regarding the use of the article) the teacher trainees are still evidencing the same difficulty as the pupils (both scored a moderate 62% only). Equally, those at the university may be faced with difficulties too. Putatively, ESP students can do with little grammatical knowledge, since the focus is primarily on functions. However, this can only be partly true, since it has been suggested that there is a minimum low level of linguistic competence as a prerequisite for adequate communicative competence to take place (Rea 1986; Politzer and McGroarty 1983; Ramirez 1984). According to Givon (1984), in discourse, the most urgent task is to establish the 'topic' (i.e. what it is that we are talking about) and then make a 'comment' about it. If L<sub>2</sub> learners are still grappling with aspects of noun reference (or the articles), which provides a means for establishing the discourse referents (i.e. topic/comment), then their communicative competence still remains in a shaky state. No one could reasonably take comfort at it, let alone the ESP teachers. It may be necessary to conduct special teaching programmes targeted at improving the learners' linguistic awareness in this area (i.e. focus on declarative knowledge, Kasper and Faerch 1983).

As far as the more communicatively oriented measures are concerned (Table 8.13- 14), we do observe a steady progress across levels ( $p < .01$  or beyond). But there still is considerable room for improvement too. According to Cook (1979:177) the style complexity measures for the native speaker's texts (we applied Cook's clause analytical technique): 1) a simple style would be characterised by measures between 1.00 and 1.40; b) a medium style 1.40 and 1.70; and 3) a complex style would involve 1.70 and above. Our results (Table 8.13) show that the best of our students score no more than 1.68 (i.e. well within the medium range of ability). Nonetheless, these results are in agreement with Eskstrand (1975) and Hatch (1983) who have argued that the

numbers of years of  $L_2$  exposure lead to greater success. According to Hatch, this may be true only as far as the overall communicative ability is concerned; but not necessarily true when grammatical or phonological accuracy are the norm.

### 8.8. Summary

In this chapter we have examined the complex relationship that is assumed to exist between the learner factors and  $L_2$  proficiency. Owing to the limited size of our sample, the results presented here are best regarded as exploratory in kind. Still they reveal a number of facts found to be true for other  $L_2$  learners, in keeping with language learning and language testing research theory. First of all, there are difficulties in defining, manipulating and quantifying variables such as attitudes and motivation, mainly because the results are based on learner self-reported questionnaires. It has been suggested that such elicitation techniques may affect the validity of the constructs under investigation (Oller and Perkins 1980; Murakami 1980). Using the factor analysis procedure it was possible to identify two separate factors for the learner variables: one 'attitudinal', the other representing 'motivation intensity towards instruction'.

Secondly, there is the difficulty of establishing the influence of the learner factors on  $L_2$  proficiency. We find that the 'attitudinal variables' do not correlate with  $L_2$  proficiency, whereas the 'motivation intensity towards instruction' does correlate with communication-based  $L_2$  measures. Finally, since our learners do not evidence social and or psychological distance, it has not been possible to test Schumann's (1978) Pidginization hypothesis. We may speculate that for many  $L_2$  learners, the amount of input, length of exposure and earlier success are the best predictors of  $L_2$  achievement. Thus social and psychological

distance may be the result rather than the cause of individual's learning experience.

## CHAPTER 9

## SUMMARY, SIGNIFICANCE OF THE STUDY AND CONCLUSION

## 9.1. Introduction

The present study has focused on the following aspects of the learner English (i.e. Interlanguage) of Zairean students: first, the use of definite and indefinite reference; secondly, the interpretation of infinitival complements and the role played by the Universal Topicality hierarchies in the identification of the correct Subject NP of the subordinate clause; thirdly, how to establish a reliable and easily applicable index of  $L_2$  global development as well as an account of the contextual variability in IL performance; fourthly, the role of individual learner factors in determining rate of  $L_2$  development and eventual success in  $L_2$  learning.

In the preceding chapters the study shows that the language of the language learner can be described in terms of a separate system (i.e. different from both the native and target languages, yet sharing some features with both) and that this linguistic system is highly variable and systematic (Nemser 1971, Selinker 1972, Richards 1973). Variability has traditionally been dismissed by those linguists whose stated aim is the discovery of an invariant structure of the linguistic knowledge of the ideal speaker-hearer of a homogeneous speech community (Chomsky, 1965). This non-procedural approach to the study of language phenomena assumes that language is a static entity; and, thus, fails to take into account the fact that grammars are intrinsically a manifestation of change. The latter, in turn, is inextricably linked to systematicity. Interlanguage grammars too exhibit such changes, and perhaps do so more rapidly than most other language phenomena within the scope of linguistics. By studying Interlanguages in terms of a dynamic and organic

system, L<sub>2</sub> researchers have been able to capture the kind of systematic variability and change inherent in developing grammars (Bailey 1973; Bickerton 1973, 1975; Huebner 1983; Meisel, Clahsen and Pienemann 1981). To a certain extent, these procedural approaches to the study of L<sub>2</sub> acquisition have been at the vanguard of research into learner language varieties over the last decade or so. So far, this line of inquiry has been instrumental in revealing a number of facts which have great significance for theories of language acquisition, classroom-oriented research and educational policy. Before exploring the significance of the findings of this study, I shall first summarize its main features and the hypotheses that it generates.

## 9.2. Features of the Current Study

This study is a cross-sectional investigation of learner English involving adult speakers of French as an L<sub>2</sub> and two or more Bantu languages, in a formal setting. The study is based on elicited performance data from learners at high school, university and college levels. A longitudinal design is arguably better suited to reflect the underlying theoretical construct of a Wave Model of language change (Bailey 1973). However, a cross-sectional design is often the only realistic option in most educational settings. The present study focuses on variability in terms of linguistic contexts. It clearly shows that IL users gradually and systematically create new linguistic contexts, such that the existence of one linguistic form presupposes the presence of one or more forms (Ellis 1986). Our intensive micro-analysis of the use of reference shows that the presence of the form of the definite article *the* does not necessarily mean that the learner has mastered its function, viz. to identify the appropriate topic vs. comment in discourse. Additionally, our analysis of the *tell/ask/promise* sentences indicates how both syntactic and semantic information and other knowledge affect the learner's understanding of these sentences. Perhaps the most

important feature of our analysis of the learners' performance is that both 'errors' and 'non-errors' as well as possible 'avoidance' of forms are accounted for in considerable detail.

Traditionally, both contrastive analysis and error analysis examine only deviations from the target norm. The order of acquisition approach, in particular, looked at IL forms in Obligatory contexts in the target language; thus leaving aside those forms which are apparently correct, but involve inappropriate uses in text or discourse context. This study emphasizes that both the 'erroneous' and 'non-erroneous' part of the learner's output are worth examining if we want to reach a better understanding of the learning process.

Another aspect of variability is that induced by learner factors. This study has investigated the effect of certain individual learner factors (e.g. attitude, motivation, length of L<sub>2</sub> exposure) on IL variability. Social and psychological factors like these are known to have some influence on L<sub>2</sub> proficiency. Let us now turn to the significance of these findings for theoretical and pedagogical purposes.

### **9.3. Significance of the Findings of the Present Study**

The main body of the findings of this study suggests that the knowledge of the rules underlying learner-language behaviour is affected by the knowledge of how, when, or where to use them. Since the two types of knowledge are interrelated (i.e. language competence appears to be heterogeneous), variability in learner speech must concern itself with both the linguistic environments and the situational demands (i.e. styles or tasks) made on the learner. The findings of a study of heterogeneous competence such as this has great significance for both L<sub>2</sub> acquisition research and L<sub>2</sub> pedagogy or educational policy. In the following paragraphs we will deal with these in turn.

### 9.3.1. Significance of the findings for L<sub>2</sub> Acquisition Research

The findings of the study (esp. Chapters 5, 6 and 7) demonstrate that the development of particular IL subsystems involves a gradual attempt to incorporate new and more complex contexts, or sets of form-function relationships. In order to go through the various stages, both L<sub>1</sub> and L<sub>2</sub> learners sometimes appear to regress in certain environments which had already been mastered. These temporary decrements in performance are called U-shaped curves and are usually considered to indicate a change in the principles by which the child copes with that particular linguistic task. This study also reveals that learner language behaviour is affected by different task types; i.e. depending on whether the focus is on more grammatically-oriented tasks or communicative ones.

The significance of these findings is that descriptions of the language of the language learner is a useful heuristic tool which can tell us a great deal about the conditions under which language transfer and other phenomena (such as language universals or markedness principles) operate.

More specifically, such descriptions can provide us with greater insights into the conditions of 'transferability' (Faerch and Kasper, 1987). By transferability we mean the set of conditions or criteria - linguistic, psychological, and socio-psychological which favour or discourage transfer. In recent years, L<sub>2</sub> researchers have pointed out the need to clarify the notion of transfer and its relationship with declarative knowledge (i.e. knowledge of 'what') as opposed to procedural knowledge (i.e. knowledge of 'how'). As Adjemian (1983) suggests, it would be an undue simplification to regard transferred L<sub>1</sub> elements as homogeneous and invariable: learners may transfer more or less of a given L<sub>1</sub> rule or item. One can thus distinguish types of

transfer in terms of how much is transferred viz. 'transfer load'. On the other hand learners not only have the option of transferring  $L_1$  knowledge to different degrees, they can also decide not to transfer at all, hence 'transfer avoidance'. The notion of transferability, then, evolves from the need to answer the wider question: 'When do learners transfer and when don't they?'. The conditions or criteria for transferability mentioned above constitute an attempt to find an answer to this complex question.

#### **9.3.1.1. Linguistic Criteria for Transferability**

The issue here is whether the typological similarities and differences between  $L_1$  and  $L_2$  can serve as a predictor for transfer. Earlier approaches to transfer, especially in the 60's, were concerned with the issue of whether typological differences between  $L_1$  and  $L_2$  were impeding or conducive to interference (James 1969, and to a certain extent Eckman 1977). More recent studies of the transferability criteria aim at providing a better understanding of the conditions under which language transfer operates (Gass 1979; Zobl 1980; Andersen 1983; Gundel and Tarone 1983).

However, from these studies of linguistic criteria for transferability, it appears that more research effort is needed to enable us to establish the predictive or explanatory power of  $L_2$  learners' transfer. This study has erred more on the explanatory side of transfer than on the predictive one.

#### **9.3.1.2. Psycholinguistic Criteria for Transferability**

The psycholinguistic approach to questions of transferability has been mostly emphasized by Kellerman (1978, 1983). According to this approach, transfer is determined by: 1) the learner's metalinguistic awareness of language distance or 'psychotypology' (Kellerman 1983); and 2) certain characteristics of  $L_1$  rules or items (psycholinguistic markedness). However, there is still a good



deal of debate among  $L_2$  researchers as to which rules or items (i.e.  $L_1$  or  $L_2$ ) should be regarded as determining the transferability of the phenomena in question. Some researchers stress that markedness is a strictly  $L_1$  related concept, independent of a particular  $L_2$  (Kellerman 1977, 1978; Jordens 1977). Others ascribe the strongest impact on transferability to the formal properties of the  $L_2$  (Zobl 1980; Andersen 1983). On the basis of our findings, it has not been possible to obtain a clear-cut answer or support for one rather than the other of these two opposing views. We can speculate however that psychological transferability depends variably on the  $L_2$  learning situation, the learner's prior linguistic and metalinguistic knowledge. Future studies are necessary to confirm or qualify the present body of knowledge in this area.

#### **9.3.1.3. Socio-psychological Criteria for Transferability**

The socio-psychological approach to transferability has great significance for  $L_2$  research, especially in the wake of increasing interest among researchers in the sociolinguistic or socio-cultural approaches to learner language varieties of e.g. Indian English (Kachru 1982); Nigerian-English (Banjo 1971); African English (Criper 1971; Bokamba 1982); Philippino English (Platt, Weber and Ho 1984). Essentially the socio-psychological criteria for transferability involve the issue whether it is possible to predict which form or IL varieties (or styles) will be activated in which contexts of use. A cognitively based answer to this question would be that the more a situation allows learners to 'monitor' their performance (Krashen 1981, 1982) the closer to the  $L_2$  norm (or formal style) the variety will be. Thus a variety activated in a writing situation (or task) will be more correct than the variety used in face-to-face interaction. Any of the possible ranges of styles existing between these two extremes of the Capability Continuum (Tarone 1983) is worth investigating.

What is lacking in a cognitive approach as referred to here is the sociolinguistic and or socio-cultural dimension of linguistic variation, which is usually the focus in most institutionalized varieties (i.e. 'new Englishes', Kachru 1983). Giles and Smith (1979) have addressed this aspect of variability by advocating what is known as the Accommodation theory. According to this theory, learners do not always adopt 'convergent behaviour'; i.e. approximate to the speech style of the (native speaker) interlocutor. There are at least three main reasons why learners 'diverge' from the norm, i.e. by increasing the amount of  $L_1$  specific features in the IL performance: 1) group solidarity, 2) foreigner role, and 3) marking the origin (Janicki 1982; Ervin Tripp 1969; Ryan 1983; Beebe and Zuengler 1983).

Group solidarity as formulated by Ryan (1983) is critical in understanding why low prestige varieties persist within individuals. In certain types of communication, low prestige language varieties may be interpreted as IL varieties with high transfer load (e.g. involving code-switching and borrowing). Cases of group solidarity leading to linguistically divergent behaviour involve ethnic minority groups marking their group-membership by preserving features of their  $L_1$ , when using the dominant  $L_2$ . Ryan (1983) points out that some native speakers of Welsh, for instance, sometimes broaden their Welsh accent in English when they believe the value of their native Welsh language is being questioned.

On the other hand, 'Foreigner role' linguistically divergent behaviour (i.e. increasing  $L_1$  transfer) seems to mark not so much group solidarity as to stress the non-membership in  $L_2$  speech community; thereby speakers of the interference or IL variety protect themselves from being assessed on the basis of native-speaker norms and expectations. Both Ervin-Tripp (1969) and Ryan

(1983) argue that L<sub>2</sub> learners often protect themselves from the consequences of sociolinguistic mistakes by preserving non-native features in their speech.

Additionally, a rather peculiar but significant case of L<sub>1</sub> transfer within a socio-psychological perspective has been known as 'Marking the origin of commodities' by preserving L<sub>1</sub> features (Faerch and Kasper 1987). 'Commodities', in the broadest possible sense of the term, include anything from purely commercial ones (e.g. clothes) to written commodities (e.g. books, newspapers, etc). In such cases, the terms used to refer to such commodities (or, in the case of written products, the commodities themselves) may exhibit linguistic features of the original culture. Faerch and Kasper (1987) point out that marking the origin of 'commodities' may have purely commercial motivations. They give the example of clothes used for jogging purposes, which are being marketed in non-English speaking countries as 'jogging X' (X belonging to the target language). Thus, in Germany today, one buys *Joggingschuhe* and *Jogginganzuege*, commodities which a few years ago were being referred to as *Turnschuhe* and *Traininganzuege*.

Beebe (1983) points out that authors of Chinese cookery books in the U.S.A., occasionally, add a flavour to the text by using Chinese-specific linguistic features. Similar examples of 'transfer', combining both foreigner role and marking the origin of 'commodities' are widespread in the learner-language varieties, i.e. what has been referred to as 'New Englishes' (Platt et al 1984). Magura (1984: 19) cites some examples of transfer whereby terms of L<sub>1</sub> meanings are transferred to an L<sub>2</sub> (English) context in his analysis of South African and Zimbabwean English. Thus, in African English, terms such as *rainmaker*, *medicine-man*, *magic-workers* and *turn-boys* fall under this category of transfer combining both foreigner role and marking the origin of 'commodities'.

In this case of *medicine-man*, for example, what African / Zimbabwean English does is transfer the Bantu word *n'anga* to a lexical item in native English. Similar cases of approaching the study of learner-English from its sociological use include Mazrui's (1975) analysis of the English language from the point of view of its political sociology. Mazrui (1975: 13) seems however, more concerned with the 'foreigner role' dimension of accommodation theory than marking the origin, since he argues that English in Africa should be not only 'de-anglicized' or 'de-racialized' but that it should of necessity be 'africanized' (i.e. adopting a linguistically divergent behaviour to mark non-membership in L<sub>2</sub> speech community). Other studies, however, seem to be more concerned with marking the origin (i.e. 'commodities' or items originating in L<sub>1</sub> culture are being transposed to a different L<sub>2</sub> English culture). Such 'New Englishes' varieties include among others, Indian English (Kachru 1965, 1982); Ghanaian English (Criper 1971; Sey 1973), Nigerian English (Banjo 1971; Bamgbōse 1971; Salami 1968; Adegoke 1969; Jibril 1982; Kirk-Greene 1971), Liberian English ( Hancock 1974; Angogo and Hancock 1980), Kenyan English (Zuengler 1982), African English (Bokamba 1982) and Zimbabwean English (Magura 1984).

It is highly misleading to assume, as some of these studies do, that each of these varieties constitute a new English language, say which could be regarded as a full fledged linguistic system, possibly only distantly related to standard English -- simply because the variety evidences certain 'ways of speaking which are readily identifiable as African' (Angogo and Hancock 1980: 73). Equally, it is misguided to suggest that these varieties cannot be subsumed under current approaches to second language acquisition and learning (Magura 1984: 11). In my view, these varieties fall into the category of learner English, in view of their heavy reliance on transfer as a production procedure (which is sensitive to the sociolinguistic and psychological constraints referred to above)

and as part of the cognitive process whereby  $L_1$  declarative knowledge shapes the interim grammar (Selinker and Lamendella 1978). Their distinctiveness is mostly functional, rather than formal (Halliday 1973, 1978); of course the rather liberal uses of the term Interlanguage in the 70's, either as a theory, an hypothesis, or a model may have unnecessarily attracted both undue reverence and unwarranted criticism. However, what the IL approach is concerned with is basically to establish the relationship between knowing and using linguistic forms, i.e. variability (Bialystok 1981). Another basic assumption underlying IL is that it is systematic (i.e. rule-governed combination of separate knowledge sources, one of which is the  $L_1$  system and another, the developing  $L_2$  based system (Adjemain 1976)). It is called ' $L_2$  based' since it is assumed to be composed both of  $L_2$  rules in complete conformity with the target norm and also idiosyncratic or immature versions of  $L_2$  rules. Both variability and systematicity as referred to here constitute the fundamental characteristics of learner language, including so-called New Englishes and other interference varieties.

The present study points out that both  $L_2$  acquisition theory and research would greatly benefit from the researchers' recognition of the crucial role of transfer as a production procedure, especially the conditions of transferability of  $L_1$  rules and items in certain communicative situations. It goes without saying that other processes and strategies such as simplification, inferencing and overgeneralization are also important features of learner language, and thus are worth investigating too.

### **9.3.2. Significance of the Findings for $L_2$ Pedagogy and Educational Policy**

In the preceding section we were concerned with showing that the variability in learner speech has great significance for  $L_2$  acquisition theory and

research. In this section we argue that the study of heterogeneous competence is relevant to L<sub>2</sub> pedagogy as well as educational policy.

### 9.3.2.1. Significance of the Findings for L<sub>2</sub> Pedagogy

Learner varieties must account for discrepancies between what has been taught and what is used and among the things which are used in different situations (Swain 1981; Tarone, Frauenfelder and Selinker 1976). According to Bialystok (1981), rather than asking 'How often does the learner produce the correct form?', the question must be formulated as 'Under what circumstances does the learner produce the correct form?' In order to answer such a question, instructional programmes must define aims in terms of conditions under which the learner will be using the target language and provide training appropriate to those specific needs. This means that language mastery must be conceived of in quantitative as well as qualitative terms, which involves the contextual and situational demands made on the learner.

In most EFL situations, both teachers and course designers usually try and deal with the constraints imposed on the learner by various language situations since these involve variable control of the IL system. As far as the Zairean educational system is concerned, the instructional programmes ought to be adjusted to integrate all the four traditional skills. In practical terms, however, reading skills, and possibly writing could be emphasized at the more advanced or tertiary level, whereas listening and speaking might be given greater attention at the secondary level (Rivers 1981). The rationale behind such a difference in emphasis between the tertiary and secondary levels is that, although the teaching of the four traditional skills together is desirable, in realistic terms, it may prove an unattainable aim to achieve substantial gains in all four skills in instructional programmes such as e.g. the ESP courses at the

university level, for which the maximum number of classroom teaching is no more than 30 hours per academic year. Since reading in English seems to be the most obvious need or skill required of these students, it (reading) has been largely recognized as the most relevant activity to focus on at the tertiary level. At the secondary level the guidelines in relation to skills separation or integration are less clear. Indeed there is a feeling that, in official statements, no provisions on this matter are envisaged to help prepare the pupils to cope with reading outside school or, later on, at the university.

In this respect, the instructional programmes or the curriculum itself should be more specific, but also flexible enough in terms of what is desirable and what is or can be actually achieved. It has been taken for granted that as a result of the ESP courses at the tertiary level, the students' reading will improve; and that, in order to improve on their reading abilities, no training in the knowledge of grammar is required. However, these assumptions need not be borne out for the following reasons. First, the amount of teaching (30 hours per year) is too small and should be increased, say up to 3 to 4 hours per week. Secondly, as shown in this study, an early learning plateau seems to have set in regarding certain areas of English syntax (e.g. the use of reference in Chapter 4 or tense and aspect in Chapter 6). Therefore some teaching of grammar may be necessary to tackle such learning problems (see White 1987).

#### **9.3.2.2. Significance of the Findings for Educational Policy**

As regards the wider framework of educational policy vis-a-vis the teaching of reading and other skills, a few remarks may be in order. It is a depressing fact of life to have to recognise that, unlike in most literate societies, reading is not part of adult life in Zaire. One can indeed point out that a great many

Zairean students are poor readers or non-readers at all. The reason they are poor readers may be because they have not been 'taught' how to read. It may also be the case that they are poor readers because they do not read at all. I am more inclined to think that the Zairean students are poor readers mainly because they do not read enough, even though the real reason may be a combination of these two. After all, in literate societies many children are apparently not taught to read at all (Clark 1976). They acquire reading in the course of everyday family life, much as they acquire spoken language. As with speaking, they need some help from the adults around them. Moreover, just as children want to talk in order to join in what is going on around them, so are they likely to want to read if reading is part of behaviour of other people in their environment. The children are motivated to learn to read not just for utilitarian or personal reasons, such as the private enjoyment of stories, but also for social reasons (Wallace 1986).

Of course it may well be the case that in Zaire we are still faced with the more basic problem of Literacy versus Reading, since for the majority of Zaireans, reading is not an established social function. However, this is understandable since, after all, there are not so many books around and the few ones that are available are too expensive for most parents, students and other members of the public.

In other words, the people (and students) do not read simply because they have nothing to read. And yet, as Wallace (1986) suggests, the question of how the learner sees himself as a potential reader in social, educational and occupational contexts is very crucial in assessing literacy needs. As regards the teaching of reading skills, then, it is fairly clear that the teacher has a daunting task facing him. First, he has to focus on the learner-reader, and then, resolve



the problem posed by the provision of adequate and stimulating reading materials (i.e. focus on the text as part of the reading process).

It has been established by reading research that good readers make fewer 'eye-fixations', with less duration, than do poor readers (Just and Carpenter 1980). Thus a reduction in eye-fixations (i.e. probably while the brain is processing text) is matched by a rise in comprehension. On the other hand, readers use graphic, syntactic, semantic and discourse information in text during their processing. The technique known as 'miscue analysis' (Goodman 1974, 1978) has revealed that a number of basic reading strategies seem to underpin the reading process. These are: 1) prediction (i.e. what the next chunk of language will be), 2) sampling (i.e. selecting the minimum information from text consistent with the prediction); 3) confirming (i.e. testing the prediction against the sample), and 4) correction (i.e. if the prediction is not confirmed, another prediction is generated). An adequate teaching methodology for teaching reading skills must integrate both the 'eye-fixation' model and 'miscue analysis' (i.e. reading errors are analysed for their similarity to, or difference from, the words in the text, and inferences are made about the underlying process). In addition, one way of making up for the lack of reading materials would be to envisage the setting up of local community literacy schemes whereby students and or members of the community would be encouraged to take part in writing projects on issues that are of concern to them.

Perhaps in time this will overcome the kind of inertia we are witnessing today in relation to reading and promoting writing by Zaireans for Zaireans. Some incentives could be created if the learner-writers are then asked to 'judge' (i.e. read and comment on ) their peers' texts. The top 10 best texts might be considered for an award at the national, or even just at the school/

community level, so that more and more learner-readers could be encouraged to participate and become learner writers themselves. The idea behind all this is that in the Zairean context, creative educational policies are required to break into the current vicious circle outlined above; that is, the students are poor readers because they are non-readers, they are non-readers because they have practically nothing to read, and so they are poor readers. Let them create something to read and read it. Finally, teacher training in Zaire must become more and more 'learner-centred' if it is going to be able to bring about the above changes. The teacher's role must be re-assessed, and to a certain extent what is needed is teacher education rather than mere teacher training. The teacher in this context will no longer be the authoritarian master but an authoritative and disciplined learning assistant (A.D.L.A.). By authoritative we mean that he will strive to become more knowledgeable of his subject matter so that he can speak about it with confidence and authority. By disciplined we mean that the teacher (i.e. A.D.L.A.) will refrain from controlling and interfering with the learning process; say, by imposing his views or spoon feeding his disciples through the outdated practice of lecture notes (known as 'Le syllabus du cours'). With such notes provided, all the students have to do is cram and then regurgitate the teacher's ideas in the form and shape of written papers at the end of the session, without understanding or integrating their content into their (students') natural or contextual environment.

Essentially, then, parrot-learning habits among the Zairean students can be seen as a direct consequence of the authoritarian teacher centred approach still in common use in teacher training colleges and other educational establishments in the country. This approach should be abandoned since it is inefficient and can stifle the learner's inventiveness or integration of knowledge. We believe that literacy schemes such as those proposed above (i.e.

integrating both reading and writing projects) may help in devolving the control of the learning process away from the teacher to the learner. Similar schemes have been conducted with encouraging results in Great Britain and elsewhere (Wallace 1986). These include, for instance, 1) the Mother tongue and English Teaching Project (Bradford University, England), 2) The friends' Centre in Brighton (England) and 3) The Gambian Literacy Scheme (Gambia, West Africa).

#### **9.4. Future Directions for Research**

##### **9.4.1. Future Research in the Area of Reference in SLA**

The present study shows that, in relation to the use of definite reference (Chapter 5), the learner is involved in making a subtle distinction between referents which are linguistically recoverable (i.e. uses of *the*) and those non-linguistically or contextually present (i.e. uses of *the*<sub>2</sub> or the case of entailment discussed by Karttunen 1968). Thus, while the use of indefinite reference (i.e. *a, an*) poses little difficulty, the distinction between the form and function of the definite articles (i.e. *the* and *the*<sub>2</sub> remains a problem and fruitful area for research). Meanwhile it would be interesting to examine how change in definite reference affects other aspects of reference such as anaphora (i.e. pronominal reference) and vice-versa. Since interlanguage is an integrated system, what is the nature of other areas of the IL at various points in time and how do they change? Future studies will have to exploit or incorporate such notions as mutual knowledge and topic vs. comment in their investigations of learner language behaviour.

##### **9.4.2. Future Research on the Role of Language Universals**

The study also demonstrates that syntax does play an important role in the comprehension of complex sentences. In Chapter 6, the study takes up the issue that universals of language (Gass 1979, 1984) have a definite influence on

the interpretation of the *tell / ask / promise* sentences. While agreeing with this suggestion (i.e. the universal topicality hierarchies play a variable role in the identification of the correct subject NP of the infinitival complements), we do not find that level of proficiency constitutes a major predictor for determining the correct subject of the subordinate clause. This is indeed in disagreement with Chomsky (1969, 1972) some of whose findings were that the older the children, the more likely they were able to identify the missing subject NP. Future studies in this area may have to take into account the fact that L<sub>2</sub> learners are cognitively better equipped to deal with some of the complexities involved in dealing with these complex structures. Adult L<sub>2</sub> learners can rely on both the principles of universal grammar and their prior linguistic knowledge of L<sub>1</sub> and L<sub>2</sub>, or on knowledge of the world to infer the resolution of the conflict pointed out in this study. Additionally, future studies will have to try to establish more accurately the strength with which these universals operate, not only at sentence level but also at text and discourse levels; thus, the studies may be able to determine whether these universal hierarchies are real both linguistically and psycholinguistically.

#### **9.4.3. Future Research on L<sub>2</sub> Development Index**

One of the unique features of this study is that it proposes a reliable and easily applicable index for assessing L<sub>2</sub> global development (Chapter 7). Using writing data, we demonstrated that an information processing approach is capable of yielding an index of development. Thus, the notion of communicational capability index (CCI) is proposed. The CCI is in fact a significant improvement on the more conventional notion of T-Unit (Hunt 1965). Some of the questions worth pursuing include: first, would the measures found to be reliable and able to discriminate among writing abilities of ESL learners be equally applicable to oral data? According to our own assumption, these

measures would be applicable to both data types. Secondly, would these measures (i.e. of clause embedding weight and grammatical accuracy) increase over time? This study shows that both measures grew in keeping with the level of proficiency. Thirdly, is it not possible that the proposed measures are affected by the learner's  $L_1$  rules and items (as discussed in 9.3.1.). On the basis of our findings, we can only speculate that the CCI could be immune to interlingual influences. The reason for this line of thinking is that our results show a high correlation with those obtained by Cook (1979), who applied the same clause embedding analytical technique to discriminate between simple and complex texts written by native speakers. In other words, this measure is insensitive to interlingual influence whereas it does tap intra-lingual differences. Nonetheless, the CCI need not be entirely immune of interlingual influences since it takes into account grammatical accuracy (i.e. % of error-free information blocks) which may be due to 'avoidance of transfer' by certain  $L_1$  speakers (Schachter 1974). Clearly, further studies will have to find out the relative importance of the  $L_1$  influence on these measures. Finally, from a pedagogic viewpoint, further studies could address the issue of whether learners can be taught to write more sophisticated texts than their linguistic abilities permit.

#### **9.4.4. Future Directions for Research on the Learner Factors**

This study also deals with the influence of certain individual learner factors on  $L_2$  development (see Chapter 8). It has been pointed out (Oller & Perkins, 1980) that the findings of surveys investigating attitudinal variables through self-reported data must be brought into question, on grounds that the questionnaires may well constitute unintentional language tests rather than tap the intended constructs. In spite of all the problems involved in such studies, we think the latter should be pursued. Obviously the researchers will have to

take great care in devising their experimental design to minimize the problem created by the use of unclear or ambiguous language or items. Both sample size and content should be representative of the intended area or construct, and the quantitative or computational methods and procedures used should be chosen with great care. Future studies will also need to address the wider issue concerning the nature of L<sub>2</sub> language competence and its relationship to school/academic achievement. For instance, to what extent is Cummins (1983) right in suggesting that the language competence required for L<sub>2</sub> cognitive/academic tasks differs qualitatively (i.e. in the degree of cognitive effort and contextual support) from that required for everyday face-to-face communication? Other more common issues of interest might be: 1) the usefulness of distinguishing between the instrumental and integrative orientation index; 2) whether finely-tuned input might be more conducive to L<sub>2</sub> development than the roughly-tuned type, or 3) whether the amount (rather than the type) of input is what matters most.

## 9.5. Conclusion

Cherchons comme cherchent ceux qui doivent trouver, et trouvons comme trouvent ceux qui doivent chercher encore parce qu'il est écrit: celui qui est au terme ne fait que commencer [Saint Augustin]

This study has presented an analysis of learner-language behaviour in a formal instructional setting. The limitations of the study are several: first, the data on which the analysis is based was obtained, through elicitation tasks only owing to the rigid timetable and current organisation in education establishments involved. We can speculate that spontaneous speech data would reveal many more interesting features of language use in discourse. Secondly, intuitional data will have to be considered in future studies, e.g. in order to detect IL development with regards to the so-called zero article (i.e. generic reference)

as well as the interpretation of certain types of complex sentences. Thirdly, our analysis has focused on variability in relation to only a limited number of IL features; and yet IL is regarded as a dynamic and organic system which means that change in one area often affects the learner's behaviour in others.

In spite of all these limitations, our study has great relevance to SLA theory, classroom research and other activities (at both national and classroom levels) that are intended to promote development and learning. We deliberately avoided the use of such catch-words as 'Applications' or 'Implications' of the findings etc. The reason for this is that these terms are misleading since they convey the impression that SLA research is about 'applying' theories from other fields (especially theoretical linguistics) or generating 'definite prescriptions' for teaching. Instead we have preferred the terms 'significance' or 'relevance' of the study to theory and practice in the classroom. This is more realistic because research in our field is best conceived of as an effort to discover relevant facts which can: 1) show how the learner's  $L_1$  knowledge and principles of universal grammar, as well as individual learner factors, do influence the development of his  $L_2$  language competence; and 2) justify teaching practice and lead to new or improved ways of doing things in a principled and research inspired manner. Because there is not a direct relationship between the linguistic theories and actual classroom learning and teaching, applications and implications often end up in desperation when the predicted outcomes do not materialize. What is needed, therefore, is a greater understanding of SLA as a process, i.e. through more and more micro-analyses of different areas of individuals' interlanguages, from a variety of  $L_1$  backgrounds and over a long period of time. This study has been only a modest contribution towards this distant end. It is hoped that many more will follow.

#### BIBLIOGRAPHY

- Adams, M. 1974. SLA in Children: A Study in experimental methods: Observation of spontaneous speech and controlled production tests. Unpublished M.A. thesis. UCLA.
- Adegoke, D. B. 1969. Standard English. *Insight*. 24 : 21-23.
- Adjemian, C. 1976. On the Nature of Interlanguage Systems. *Language Learning*. 26. 297-320.
- 1983. The Transferability of Lexical Properties. In S. Gass and L. Selinker (eds).
- Alderson, J. and A. Urquhart. (eds) 1984. *Reading in a Foreign Language*. London: Longman.
- Allwright, R. L. 1975. Problems in the Study of the Language Teacher's Treatment of Learner Error. In M. Burt and H. Dulay (eds). *On TESOL*. 1975. Washington, D.C.: TESOL.
- 1978. An Attempt to Model the Role of Cognitions in Language Learning, in *Stuttgart Congress Papers*. Volume on Pragmalinguistics. Stuttgart: Hochschulverlag.
- 1980. Turns, Topics and Tasks: Patterns of Participation in Language Learning and Teaching. In Larsen-Freeman, D. (ed).
- 1984. The Analysis of Discourse in Interlanguage Studies: The Pedagogical Evidence. Paper presented at the Seminar on Interlanguage in honour of S.P. Corder. Ms. Edinburgh University Dept. of Linguistics.
- Andersen, R. W. 1977. The Impoverished State of Cross-sectional Morpheme Acquisition / Accuracy Methodology. In C. Henning (ed). *Proceedings of the Los Angeles SLA Research Forum*. Los Angeles: University of California Press.
- 1978. A Theoretical model of Second Language Learning. *Language Learning*. 28, pp. 69-84.
- 1979. Expanding Schumann's Pidginization Hypothesis. *Language Learning*. 29, pp. 105-120.
- 1981. (ed) *New Dimensions in Second Language Acquisition Research*.
- 1983. (ed). *Pidginization and Creolization as Language Acquisition*. Rowley, Mass.: Newbury House.
- 1984. (ed) *Second Languages: A Cross-linguistic Perspective*. Rowley, Mass.: Newbury Press.
- Anderson, J. M. 1971. *The Grammar of Case: Towards a Localist Theory*.



Cambridge Studies in Linguistics. 4. Cambridge University Press.

- Angogo, R. and I. Hancock. 1980. English in Africa: Emerging Standards or Diverging Regionalisms? *English Worldwide: Journal of Varieties of English*. 1, 1. 67-96.
- Anisfeld, M. and W. E. Lambert. 1961. Social and Psychological Variables in Learning Hebrew. *Journal of Abnormal and Social Psychology*. 63. 524-528.
- Austin, J. L. 1962. *How to do things with words*. Oxford: Clarendon.
- Baars, B. 1978. The Competing Plans Hypothesis: An Heuristic Viewpoint on the Causes of Errors in Speech. In Dechert and Raupach (eds).
- Bailey, B. 1971. Jamaican Creole: Can Dialect Boundaries be Defined? In D. Hymes, ed.
- Bailey, C. 1973. *Variation and Linguistic Theory*. Arlington, VA.: Center for Applied Linguistics.
- Bailey, K. 1980. An Introspective Analysis of an Individual's Language Learning Experience. In R. Scarella and S. Krashen (eds). *Research in SLA*. Rowley, Mass.: Newbury House.
- 1983. Competitiveness and Anxiety in Adult Second Language Learning: Looking at and through the Diary Studies. In H. Seliger and M. Long (eds).
- Bailey, N. C. Madden and S. Krashen. 1974. Is There a 'Natural Sequence' in Adult Second Language Learning? *Language Learning* 24, 235-243.
- Baker, C. L. 1966. Definiteness and Indefiniteness in English. Unpublished M. A. Thesis. University of Illinois.
- Bamgbose, A. 1971. The English Language in Nigeria. In J. Spencer (ed).
- Banjo, A. 1971. Standards of Correctness in Nigerian English. *West African Journal of Education*. 15, 2. 123-127.
- Bauman, R. and J. Sherzer. 1974 (eds). *Explorations in the Ethnography of Speaking*. Cambridge: Cambridge University Press.
- Beebe, L. M. and J. Zuengler. 1983. Accommodation Theory: An Explanation for Style Shifting in Second Language Dialects. In N. Woltson and E. Judd (eds) *Sociolinguistics and SLA*. Rowley, Mass.: Newbury House.
- Beebe, L. 1980. Sociolinguistic Variation and Style Shifting in Second Language Acquisition. *Language Learning* 30, 433-448.
- 1983. Risk-taking and the Language Learner. In H. Seliger and M. Long (eds).
- Bellugi, U. and R. Brown. 1968. (eds). *The Acquisition of Language. Monograph of*

*the Society for Research in Child Development*. 29. 1.

- Bellugi, U. 1967. The Acquisition of the system of negation in children's speech. Unpublished doctoral dissertation. Harvard University.
- Berktau, J. 1974. An Analysis of English Learner Speech. *Language Learning*, 24, 279-286.
- Bever, T. G. 1972. Perception, Thought and Language. In *Language Comprehension and the Acquisition of Knowledge*. ed. R.O. Freedle and J.B. Carroll. Washington, D.C.: Winston & Sons.
- Bialystok, E. and E. B. Ryan. 1985. A Metacognitive Framework for the Development of First and Second Language Skills. In D.L. Forrest-Pressley (ed). *Metacognition, Cognition and Human Performance*. New York: Academic Press.
- Bialystok, E. and M. Sharwood Smith. 1985. An Evaluation of the Construct of Second-Language Acquisition. *Applied Linguistics*, vol 6, no 2.
- Bialystok, E. 1982. On the Relationship between Knowing and Using forms. *Applied Linguistics* III. 181-206.
- 1983. Learning Language Form and Language Function. *Interlanguage Studies Bulletin*. Vol. 7. 54-70.
- Bickerton, D. and C. Odo. 1976. Change and Variation in Hawaiian English, I: General Phonology and Pidgin Syntax. Honolulu: Social Sciences and Linguistics Institute: University of Hawaii Press.
- Bickerton, D. 1971. Inherent Variability and Variable Rules. *Foundations of Language*. 7. 457-492.
- 1973. The Structure of Polylectal Grammars. *Georgetown Monographs* 25.
- 1975. *Dynamics of a Creole System*. London: Cambridge University Press.
- 1981a. *Roots of Language*. Ann Arbor: Karoma Publishers.
- 1981b. Discussion. In Andersen, R. W. (ed). *New Dimensions in Second Language Acquisition Research*. Rowley, Mass: Newbury House.
- Bloom, L. , P. Lightbown, and L. Hood. 1975. Structure and Variation in Child Language. *Monographs of the Society for Research in Child Development*. 40. 2. Chicago: University of Chicago Press.
- BMDP. 1983. Biomedical Computer Programs. In W.J. Dixon and M.B. Brown (eds). Los Angeles: University of California Press.
- Bokamba, E. G. 1982. The Africanization of English. In B. Kachru ed. *The Other Tongue: English Across Cultures*. Urbana: University of Illinois Press.
- Bolinger, D. 1980. Syntactic Diffusion and the Indefinite Article. MS Indiana University Linguistics Club.

- Borland, H. 1983. The Acquisition of Some Features of English Syntax by four groups of Adolescent Migrants to Australia. Unpublished Ph. D. thesis. University of Edinburgh.
- Brown, G. and G. Yule. 1983. *Discourse Analysis*. London: Cambridge University Press.
- Brown, H. 1980. The Optimal Distance Model of Second Language Acquisition. *TESOL quarterly*. 14. 157-164.
- Brown, R. 1973. *A First Language: The Early Stages*. Cambridge, Mass: Harvard University Press.
- Brown, R., C. Cazden and U. Bellugi-Klima. 1968. The Child's Grammar from I to III. In J. P. Hill (ed). *Minnesota Symposia on Child Language*. Minneapolis: University of Minnesota Press.
- Brown, R. and C. Fraser. 1964. The Acquisition of Syntax. In Bellugi and Brown (eds).
- Brown, R. and C. Hanlon. 1970. Derivational Complexity and Order of Acquisition in Child Speech. In J. R. Hayes (ed). *Cognition and the Development of Language*. New York: Wiley.
- Burstall, C. 1979. Primary French in the Balance. Pride, J.B. (Ed) *Sociolinguistic Aspects of Language Learning and Teaching*. Oxford: Oxford University Press.
- Burt, M., and H. Dulay. 1977. Remarks on Creativity in Language Acquisition. In M. Burt, H. Dulay and M. Finocchiaro (eds). *Viewpoints on English as a Second Language*. New York: Regent Press.
- Byrne, D. 1967. *Progressive Picture Compositions: Pupils' Book*. London: Longman.
- Canale, M. 1983. On Some Dimensions of Language Proficiency. In J. W. Oller (ed).
- Candlin, C. 1981. *The Communicative Teaching of English*. Longman.
- Carrell, P. L. 1984. Inferencing in ESL: Presuppositions and Implications of Factive and Implicative Predicates. *Language Learning*, 34, 1.
- Carroll, B. J. 1980. *Testing Communicative Performance*. Oxford: Pergamon Press.
- Carroll, J. B. 1967. Foreign Language Proficiency Levels attained by Language Majors near Graduation from College. *Foreign Language Annals*. 1. 131-151.
- 1983. Psychometric Theory and Language Testing. In J. W. Oller (ed).
- Cattell, R. B. 1966. The Scree test for the Number of Factors. *Multivariate Behavioural Research*. 1. 245-276.
- Cazden, C., H. Cancino, E. Rosansky and J. Schumann. 1975. Second Language

Acquisition Sequences in Children, Adolescents and Adults. Final Report, U.S. Dept. of Health, Education and Welfare.

Cedergren, H. and D. Sankoff. 1973. Variable Rules: Performance as a Statistical Reflection of Competence. Mimeo. (also available in *Language* 50. 333-355, 1974).

Celce-Murcia, M. 1980. Contextual Analysis of English: Application to TESL. In Larsen-Freeman, D. (ed).

Chafe, W. 1970. *Meaning and the Structure of Language*. Chicago: University of Chicago Press.

----- 1979. The Flow of Thought and the Flow of Language. In Givon ed.

Chomsky, C. 1969. *The Acquisition of Syntax in Children from 5-10*. Cambridge, Mass.: MIT Press.

----- 1972. stages in Language Development and Reading Exposure. *Harvard Educational Review*. 42. 1-33.

Chomsky, N. 1957. *Syntactic Structures*. London: Mouton.

----- 1965. *Aspects of a Theory of Syntax*. Cambridge, Mass.: MIT Press.

----- 1975. Questions of Form and Interpretation. *Linguistic Analysis* 1. 75-109.

----- 1978. A Theory of Core-Grammar. *GLOT* 1. 7-26.

----- 1980. *Rules and Representations*. New York: Columbia University Press.

----- 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

----- 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge: MIT Press.

Chomsky, N. and H. Lasnik. 1977. Filters and Control. *Linguistic Inquiry*. 8. 425-504.

Chumbow, B. 1971. Le Francais au Cameroun Occidental Interferences Phonologiques dans l'Apprentissage du francais en Milieu Plurilingue. MS. Universite Lovanium, Kinshasa.

----- 1978. The Mother Tongue Interference Hypothesis in a Multilingual Setting. MS. Paper presented at the 5th International Congress of Applied Linguistics. Montreal, Canada.

Clark, H. and S. Haviland. 1977. Comprehension and the Given-New Contract. In R. Freedle (ed).

Clark, M. 1976. *Young Fluent Readers*. Heinemann Educational Publications.

Clements, G. and S. J. Keyser. 1983. *CV Phonology: A Generative Theory of the*

*Syllable*. Cambridge: MIT Press.

- Cohen, A. 1984. Studying Second-Language Learning Strategies: How do we get Information. *Applied Linguistics* 5, No. 2.
- Cole, P. and J. Morgan. 1975. *Syntax and Semantics: Speech Acts*. New York: Academic Press.
- Comrie, B. 1981. *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell.
- 1984. Why do linguists need Language Acquirers? In W. Rutherford (ed).
- Comrie, B. and E. Keenan. 1977. Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8. 63-99.
- Corder, S. P. 1967. The Significance of Learners' Errors'. *I.R.A.L.* 5, No. 6, 160-170.
- 1973. *Introducing Applied Linguistics*. Harmondsworth: Penguin.
- 1978. Language Continua and the Interlanguage Hypothesis. In J. C. Richards ed *Understanding Second and Foreign Language Learning. Issues and Approaches*. Rowley, Mass.: Newbury House.
- 1981a. and J. Schumann. The Pidginization Hypothesis: A Model for Second Language Acquisition. *Applied Linguistics*. 2.
- 1981b. *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- 1983. A Role for the Mother Tongue. In S. Gass and L. Selinker (eds).
- Cook, V. 1973. The Comparison of Language Development in Foreign Adults. *I.R.A.L.* XI. 13-28.
- 1975. Strategies in the Comprehension of Relative Clauses. *Language and Speech* 18. 204-212.
- 1985. Universal Grammar and Second Language Learning. *Applied Linguistics*. 6.1. 2-18.
- Cook, W. A. 1973. Verb Classification in Case Grammar. University of Michigan Papers in Linguistics (UMPIL) 1, 2. 32-42.
- 1979. *Case Grammar: Development of the Matrix Model (1970- 1978)*. Washington, D.C.: Georgetown University Press.
- Cooper, R., E. Olshtain, G. R. Tucker and M. Waterbury. 1979. The Acquisition of Complex English Structures by Adult Native Speakers of Arabic and Hebrew. *Language Learning* 29, 1.
- Criper, L. 1971. A Classification of Types of English in Ghana. *Journal of African Languages* 10, 3. 6-17.

- Crombie, W. 1985a. *Discourse and Language Learning: A Relational Approach to Syllabus Design*. Oxford: Oxford University Press.
- 1985b. *Process and Relation in Discourse and Language Learning*. Oxford: Oxford University Press.
- Crystal, D. 1966. Specification and English Tenses. *Journal of Linguistics*, 2, 1.
- Cummins, J. 1979. Linguistic Interdependence and the Educational Development of Bilingual children. *Review of Educational Research*, 49, 2. 222-251.
- 1983. Language Proficiency and Academic Achievement. In J. W. Oller (ed).
- Cziko, G. 1984. Some Problems with Empirically-based Models of Communicative Competence. *Applied Linguistics* 5, 1.
- D' Anglejan, A. and R. Tucker. 1975. The Acquisition of Complex English Structures by Adult Learners. *Language Learning*, 25, 281-296.
- Dakin, J. 1973. *The Language Laboratory and Language Learning*. Longman.
- Davies, A., C. Cripser and A. Howatt (eds). 1984. *Interlanguage: Proceedings of the Seminars in Honour of Pit Corder*. Edinburgh: Edinburgh University Press.
- Davies, E. E. 1985. Communication as a Criterion for Error Evaluation. *I.R.A.L.* 23, 1. 65-69.
- DeCamp, D. 1971. Toward a Generative Analysis of a Post-Creole Speech Continuum. In D. Hymes ed.
- De Villiers, J. P. and De Villiers. 1973. A Cross-sectional study of the Acquisition of Grammatical Morphemes in Child Speech. *Journal of Psycholinguistic Research* 2. 267-278.
- Dechert, H. W. 1983. How a Story is Done in a Second Language. In Faerch, G and Kasper, G (eds).
- Dickerson, L. 1975. The Learner's Interlanguage as a System of Variable Rules. *TESOL Quarterly*, 9, 401-408.
- Dickerson, W. 1976. The Psycholinguistic Unity of Language Learning and Language Change. *Language Learning* 26. 215-232.
- Drubig, B. 1972. Zur Analyse syntaktischer Fehlleistungen. In G. Nickel *Fehlerkunde*. Berlin: Cornelsen. 18-91.
- Dryer, M. 1980. The Positional Tendencies of Sentential Noun Phrases in Universal Grammar. *Canadian Journal of Linguistics*, 25. 123-195.
- Dulay, H. and M. Burt. 1973. Should We Teach Children Syntax? *Language Learning*, 32, 2.

- 1974. A New Perspective on the Creative Analysis Hypothesis. *Language Learning*. 24. 253-278.
- 1974. Natural Sequences in Child Second Language Acquisition. *Language Learning*. 24. 253-278.
- Duranti, A. 1979. Object Clitic Pronouns in Bantu and the Topicality Hierarchy. *Studies in African Linguistics*. 10. 31-45.
- Duskova, L. 1969. On Sources of Errors in Foreign Languages. *I.R.A.L.* 7. 11-36.
- Eckman, F. 1977. Markedness and the Contrastive Analysis Hypothesis. *Language Learning*. 27. 315-330.
- 1981. On the Naturalness of Interlanguage Phonological Rules. *Language Learning* 31. 195-216.
- 1984. Universals, Typologies and Interlanguages. In W. Rutherford (ed).
- Ekstrand, L. 1975. Age and Length of Residence as Variables Related to Adjustment of Migrant Children with Special reference to Second Language Learning. Paper.
- Elliot, D. , S. Legum and S. A. Thompson. 1969. Syntactic Variation as Linguistic Data. In R. Binnick et al. (eds). *Papers from the 5th Regional Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago Press. 52-59.
- Ellis, R. 1982. The Origins of Interlanguage. *Applied Linguistics*. 3. 207-223.
- 1984a. *Classroom Second Language Development*. Oxford: Pergamon.
- 1984b. Sources of Variability in Interlanguage. Paper presented at the Interlanguage Seminar in Honour of Pit Corder, Edinburgh.
- 1984c. Formulaic Speech in Early Classroom Second Language Development. In J. Handscombe, R. Orem, and B. Taylor (eds). *On TESOL '83: The Question of Control*. Washington, D.C.: TESOL.
- 1984d. Communication Strategies and the Evaluation of Communicative Performance. *ELT Journal*. 38, 1. 39-44.
- 1984e. Can Syntax Be Taught? *Applied Linguistics*. 5, 2.
- 1986. *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- Enkvist, N. E. 1973. Should We Count Errors or Measure Success? In J. Svartvik (ed). 16-23.
- Ervin-Tripp, S. M. 1969. Sociolinguistic Rules of Address. In J. B. Pride and J. Holmes (eds). *Sociolinguistics*. Harmondsworth: Penguin.

- 1974. Is Second Language Learning Like the First? *TESOL Quarterly*. 8, 2.
- Faerch, C. and G. Kasper. 1987. Perspectives on Language Transfer. *Applied Linguistics*. 8, 2. 111-136.
- Faerch, C. , K. Haastруп and R. Phillipson. 1984. *Learner Language and Language Learning*. Clevedon, England: Multilingual Matters Ltd.
- Faerch, D. and G. Kasper. 1984. Two Ways of Defining Communication Strategies. *Language Learning*, 34, No. 1.
- Faerch, C. and G. Kasper (eds). 1983. *Strategies in Interlanguage Communication*. London: Longman.
- 1980. Plans and Strategies in Foreign Language Learning and Communication. *Interlanguage Studies Bulletin*. Utrecht. 5. 47-118.
- Farhady, H. 1983. New Directions for ESL Proficiency Testing. In J. W. Oller (ed).
- Fasold, R. 1978. Language Variation and Linguistic Competence. In D. Sankoff (ed) 1978.
- Fathman, A. 1975. The Relationship between Age and Second Language Productive Ability. *Language Learning* 25. 245-253.
- Feenstra, H. J. and R. C. Gardner. 1968. Aptitude, Attitude and Motivation in SLA. Research Bulletin no. 101. University of Western Ontario.
- Felix, S. 1977. Bist du weiss was sein - I don't. *Linguistische Untersuchungen zum englisch deutschen Zweitsprachenerwerb unter natuerlichen Bedingungen*. Habilitation thesis. University of Kiel.
- Fillmore, C. J. 1966. Toward a Modern Theory of Case in E. Bach and R. Harms (eds). *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston.
- Frawley, W. 1981. The Complement Hierarchy, Evidence for Language Universals from L<sub>2</sub>. Paper presented at Winter SLA.
- Freed, B. 1980. Talking to Foreigners versus talking to Children: Similarities and Differences. In R. Scarcella and S. Krashen (eds) 1980. *Research in SLA*. Rowley, Mass.: Newbury House.
- Freedle, R. (ed). 1977. *Discourse Production and Comprehension*. Norwood, N.J.: Ablex Publishing.
- Friedlander, B., A. Jacobs, B. Davis, and H. Wetstone. 1972. Time-sampling analysis of infants' natural language environments in the home. *Child Development*. 43. 730-740.
- Gaies, S. J. 1976. Sentence Combining: A Technique for Assessing proficiency in a Second Language. Paper presented at the conference on Perspectives on Language. University of Louisville, Kentucky.



- 1983. Learner Feedback: An Exploratory Study of its Role in Second Language Classroom. In H. W. Seliger and M. H. Long (eds).
- Gardner, R. C. 1960. Motivational Variables in Second Language Learning. Ph. D. Dissertation. McGill University.
- Gardner, R. C. and W. E. Lambert. 1972. *Attitudes and Motivation in Second Language Learning*. Rowley, Mass.: Newbury House.
- Gardner, R. C. and E. H. Santos. 1970. Motivational Variables in Second Language Acquisition: A Philippine Investigation. *Research Bulletin no. 149*. University of Western Ontario.
- Gardner, R. C. and P. Smythe. 1975. Second Language Acquisition: A Social psychological Approach. *Research Bulletin no. 332*. University of Western Ontario.
- Gardner, R. C., P. Smythe and G. Brunet. 1977. Intensive Second Language Study: Effects on Attitudes, Motivation and French Achievement. *Language Learning* 27. 243-262
- Gardner, R. C., P. Smythe and R. Clement. 1979. Intensive Second Language Study in a Bicultural Milieu: An Investigation of Attitudes, Motivation and Language Proficiency. *Language Learning*. 29, 2. 305-320.
- Gass, S. 1979. Language Transfer and Universal Grammatical Relations. *Language Learning*. 29. 327-344.
- 1982. From Theory to Practice. *On TESOL '81*. W. Rutherford and M. Hines (eds). Washington, D.C.: TESOL.
- 1984. The Empirical Basis for the Universal Hypothesis. In I. L. Studies. Paper presented at the Seminar on Interlanguage in honour of S. P. Corder. MS. Edinburgh Univ. Ling. Dept.
- Gass, S. and J. Ard. 1984. Second Language Acquisition and the Ontology of Language Universals. In W. Rutherford (ed).
- Gass, S. and L. Selinker (eds). 1983. *Language Transfer in Language Learning*. Rowley, Mass.: Newbury House.
- Giles, H. 1973. Accent Mobility: A Model and Some data. *Anthropological Linguistics*. 15. 87-105.
- 1977. ed. *Language, Ethnicity and Intergroup Relations*. London: Academic Press.
- Giles, H., R. Bourhis and D. Taylor. 1977. Towards a Theory of Language in Ethnic Group Relations. In H. Giles (ed). 1977.
- Giles, H. and B. St. Jacques. 1979. *Language and Ethnic Relations*. Oxford: Pergamon.
- Giles, H. and P. M. Smith. 1979. Accommodation Theory: Optimal Levels of

- Convergence. In H. Giles and R. M. St. Clair (eds) *Language and Social Psychology*. Baltimore: University Press.
- Givon, T. 1976. Topic, Pronoun and Grammatical Agreement. In C. Li (ed) *Subject and Topic*. New York: Academic Press.
- 1979. ed. *Syntax and Semantics*. New York: Academic Press.
- 1982. Tense-Aspect - Modality in Creole. In Hopper, P. (ed).
- 1984. Universals of Discourse Structure and Second Language Acquisition. In W. Rutherford (ed).
- Goodman, K. S. 1974. Miscue Analysis: Theory and Reality in Reading. Paper presented at the 5th World IRAL Congress. 1974. In J. Merrit (ed). *Proceedings, New Horizons In Reading*. International Reading Association. Vol. 19. 1976.
- 1976. Reading: A Psycholinguistic Guessing Game. In H. Singer and R. B. Ruddell (eds). *Theoretical Models and Processes of Reading*. International Reading Association. 2nd edition.
- 1978. Miscue Analysis. Quoted in F. Smith. ed. Reading. Cambridge University Press.
- Gough, J. 1975. Comparative Studies in Second Language Learning. Unpublished M.A. Thesis. Los Angeles: UCLA.
- Greenberg, J. C. 1966. Some Universals of Grammar with Particular reference to the Order of Meaningful Elements. In J. Greenberg (ed). *Universals of Language*. Cambridge, Mass.: MIT Press.
- Guilford, J. P. and B. Fruchter. 1978. *Fundamental Statistics in Psychology and Education*. London. McGraw-Hill.
- Guiora, A. and W. Acton. 1979. Personality and Language: A Restatement. *Language Learning*. 29. 193-204.
- Gundel, J. and E. Tarone. 1983. Language Transfer and the Acquisition of Pronominal Anaphora. In S. Gass and L. Selinker (eds).
- Guttman, L. 1944. A Basis for Scaling Qualitative Data. *American Sociological Review*. 9. 139-150.
- Hakuta, K. 1974. Prefabricated patterns and the Emergence of Structure in Second Language Acquisition. *Language Learning*. 24. 287-297.
- 1975. Becoming Bilingual at Age Five: The Story of Uguisu. Unpub. Ph. D. thesis, Harvard University.
- 1981. Some Common Goals for Second and First Language Acquisition Research in Anderson, in Andersen, R. W. (ed).
- Halliday, M. 1973. *Explorations in the Functions of Language*. London: Edward Arnold.

- Hancock, I. 1974. Identity, Equality and Standard Language. *Florida FL Reporter*. Spring / Fall 49-52. 101-102.
- 1975. Problems in the Creation of a Standard Dialect of Romanes. *Working Papers in Sociolinguistics* 25. S.W. Educational Development Laboratory, Austin, Texas.
- Harris, D. 1969. *Testing English as a Second Language*. London: McGraw-Hill.
- Hatch, E. 1978. *Discourse Analysis and Second Language Acquisition. A Book of Readings*. Rowley, Mass.: Newbury House.
- 1978 (ed). *Second Language Acquisition*. Rowley, Mass.: Newbury House.
- 1983. Simplified Input and Second Language Acquisition. In R. Andersen (ed) *Pidginization and Creolization as Language Acquisition*. Rowley, Mass.: Newbury House.
- 1983. *Psycholinguistics. A Second Language Perspective*. Rowley, Mass.: Newbury House.
- Hatch, E. and H. Farhady. 1982. *Research Design and Statistics for Applied Linguistics*. Rowley, Mass.: Newbury House.
- Hatch, E. and M. H. Long. 1980. Discourse Analysis, What's That? In Larsen-Freeman, D. (ed).
- Hatch, E. and J. Wagner-Gough. 1976. Explaining Sequence and Variation in SLA. In H. D. Brown (ed) *Papers in Second Language Acquisition*. Ann Arbor: University of Michigan Press.
- Hawkins, J. 1983. *Word Order Universals*. New York: Academic Press.
- Hawkinson, A. and L. Hyam. 1974. Hierarchies of Natural Topic in Shona. *Studies in African Linguistics*. 5. 147-170.
- Hennig, G. 1976. A Developmental Analysis of Errors of Adult Iranian students of English. *UCLA-USC Second Language Acquisition Forum*. Los Angeles.
- Hockey, S. and I. Marriott. 1980. *Oxford Concordance Program*. Oxford University Press.
- Hopper, P. J. (ed). 1982. *Tense-Aspect: Between Semantics and Pragmatics*. Amsterdam: John Benjamins.
- Huang, IJ. 1971. A Chinese Child's Acquisition of English Syntax. Unpublished M.A. thesis. UCLA.
- Huebner, T. 1979. Order-of-Acquisition vs. Dynamic Paradigm: A Comparison of method in Interlanguage Research. *TESOL Quarterly*. 13. 21-28.
- 1983. *A Longitudinal Analysis of the Acquisition of English*. Ann Arbor: Karoma Publishers.

- 1984. System and Variability in Interlanguage Syntax. *Language Learning*, vol 35., no 2.
- 1985. System and Variability in Interlanguage Syntax. *Language Learning*, 35, 2. 141-163.
- Hughes, A. and C. Lascaratou. 1982. Competing Criteria for Error Gravity. *ELT Journal*, 36, 3. 175-182.
- Hunt, K. W. 1965. Grammatical structures written at three grade levels. NCTE Research Report. No. 3. Champaign, Ill.: National Council of Teachers of English, 1965.
- Hyltenstam, K. 1977. Implicational Patterns in Interlanguage. *Language Learning*, 27. 383-412.
- 1982. Language, Typology, Language Universals, Markedness and SLA. Paper presented at the Second European-North American Workshop of Second Language Acquisition research. Goehrde, Germany.
- Jakobovits, L. A. 1970. *Foreign Language Learning*. Rowley, Mass.: Newbury House.
- James, C. 1969. The Exculpation of Contrastive Linguistics. In G. Nickel (ed) *Papers in Contrastive Linguistics*. Cambridge: Cambridge University Press.
- 1980. *Contrastive Analysis*. London: Longman.
- Janicki, K. 1982. *The Foreigner's Language in a Sociolinguistic Perspective*. Poznan: Adam Mickiewicz University.
- Jespersen, O. 1914 (1927). *A Modern English Grammar*. Heidelberg: Carl Winter's Universitaetbuchhandlung.
- 1933. *Essentials of English Grammar*. University of Alabama Press.
- Jibril, M. 1982. Nigerian English: An Introduction. In J. B. Pride (ed). *New Englishes*. Rowley, Mass.: Newbury House.
- Johansson, S. 1973. The Identification and Evaluation of Errors in Foreign Languages: A Functional Approach. In J. Svartvik, ed. *Errata: Papers in Error Analysis*. Lund: CWK Gleerup.
- Jones, R. 1975. Probability Estimation using a Multinomial Logistic Function. *Journal of Statistical Computation and Simulation*, 3. 315-329.
- Joos, M. 1971. Hypothesis as to the Origin and Modifications of Pidgins. In D. Hymes ed.
- Jordens, P. 1977. Rules, Grammatical Intuitions and Strategies in Foreign Language Learning. *Interlanguage Studies Bulletin*. Utrecht 212. 5-76.
- 1986. Production Rules in Interlanguage Evidence from Case Errors in L<sub>2</sub>

- German. In E. Kellerman and M. Sharwood-Smith (eds). *Crosslinguistic Influence in SLA*. Oxford: Pergamon.
- Just, M. and P. Carpenter. 1980. A Theory of Reading - from Eye Fixations to Comprehension. *Psychological Review*. 87. 329-354.
- Kachru, Y. 1982. Linguistics and Written Discourse in Particular Languages: English and Hindi. *ARAL* 3. 50-77.
- 1982. Models for Non-native Englishes. In Kachru ed.
- 1983. Cross-Cultural Texts, Discourse Strategies and Discourse Interpretation. In L. Smith (ed). *Discourse across Cultures: Strategies in World Englishes*. Oxford: Pergamon Press.
- Kachru, B. 1965. The Indiannes in Indian English. *Word*. 21. 391-410.
- Kachru, B.B.(ed).1982. *The Other Tongue*. University of Illinois Press.
- Kaiser, H. F. 1958. The Varimax Criterion for Analytic Rotation in Factor Analysis. *Psychometrika*. 23. 187-200.
- Karttunen. L. 1968a. *What do Referential Indices Refer to?* . Santa Monica, California: The Rand Corporation.
- 1968b. *What makes Definite Noun Phrases Definite?* Santa Monica, California: The Rand Corporation.
- Kasper, G. 1982. Teaching-induced Aspects of Interlanguage Discourse. *Studies in Second Language Acquisition*. 4. 99-113.
- Keenan, E. 1978. Language Variation and the Logical Structure of Universal Grammar. In H. Seiler (ed). *Language Universals*. Tubinger Beitrage zur Linguistik III (Tubingen: Gunter Narr). 89-123.
- Keenan, E. and B. Comrie. 1977. Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry*. 8. 63-99.
- Kellerman, E. 1977. Towards a Characterization of the Strategy of Transfer in Second Language Learning. *Interlanguage Studies Bulletin*. Utrecht. 2. 58- 145.
- 1978. Giving Learners a Break: Native Language Intuitions as a Source of Predictions about Transferability. *Working Papers on Bilingualism*. 5. 60-92.
- 1979. Transfer and non-transfer, where are we now? *Studies in Second Language Acquisition*. 2. 37-57.
- 1983. Now You see it, now you don't. In S. Gass and L. Selinker (eds) 1983.
- Kiparsky, P. and Kiparsky, C. 1970. Fact in Bierwisch, M. and Heidolph, K. E. (eds) *Progress in Linguistics*. The Hague: Mouton.
- Kirk-Greene, A. 1971. The Influence of West African Languages on English. In

- J. Spencer (ed). *The English Language in Africa*. London: Longman.
- Klein, W. and N. Dittmar. 1979. *Developing Grammars: The Acquisition of German Syntax by Foreign Workers*. Berlin and New York: Springer.
- Krashen, S. 1973. Lateralization, Language Learning and the Critical Period: Some New Evidence. *Language Learning* 23, 1. 63-74.
- 1976. Formal and Informal Linguistic Environments in Language Acquisition and Language Learning. *TESOL Quarterly*. 10. 157-168.
- 1977. The Monitor Model for Adult Second Language Performance. In M. Burt, H. Dulay, and M. Finochiaro *Viewpoints on English as a Second Language*. New York: Regents Publishers.
- 1979. A Response to McLaughlin, "The Monitor Model": Some Methodological Considerations. *Language Learning*. 29. 151-168.
- 1981. *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon.
- 1982. *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.
- Krashen, S., C. Jones, S. Zelinski and C. Usprich. 1978. How Important is Instruction? *ELT Journal*. 32. 257-261.
- Krashen, S. and H. Seliger. 1976. The Role of Formal and Informal Linguistic Environments in Adult Second Language Learning. *International Journal of Psycholinguistics*. 3. 15-21.
- Krashen, S., H. Seliger and D. Harnett. 1974. Two Studies in Adult Second Language Learning. *Kritikon Litterarum*. 3. 220-228.
- Krashen, S., V. Sferlazza, O. Feldman and A. Fathman. 1976. Adult Performance on the SLOPE Test: More Evidence for a Natural Sequence in Adult Second Language Acquisition. *Language Learning*. 26. 145-152.
- Krashen, S. and T. Terrell. 1983. *The Natural Approach. Language Acquisition in the Classroom*. Oxford: Pergamon Press.
- Kumpf, L. 1982. An Analysis of Tense, Aspect and Modality in Interlanguage. Paper presented at TESOL, Honolulu.
- 1983. A Case Study of Temporal Reference in Interlanguage in Campbell, C. et al. (eds). *Proceedings of the Los Angeles L<sub>2</sub> Research Form*. Los Angeles, UCLA.
- Kuroda, S. Y. 1966. English Relativization and Certain Related Problems. In Reibel and Schane (eds) *Modern Studies in English*. Englewood Cliffs, N. J.: Prentice Hall.
- Labov, W. 1966. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

- 1969. Contraction, Deletion and Inherent Variability of the English Copula. *Language*. 45. 715-762.
- 1970. The Study of Language in its Social Context. *Studium Generale*. 23. 30-87.
- 1972. *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.
- 1972. The Study of Language in its Social Context, in Pride and Holmes (eds) *Sociolinguistics*. Harmondsworth, Middlesex: Penguin.
- 1980. *Locating Language in Time and Space*. New York: Academic Press.
- Labov, W. and T. Labov. 1977. L'apprentissage de la syntaxe des interrogations. *Langue Francaise*. 34. 52-80.
- Lado, R. 1957. *Linguistics Across Cultures. Applied Linguistics for Language Teachers*. University of Michigan Press.
- 1964. *Language Teaching: A Scientific Approach*. New York: McGraw-Hill.
- Lambert, W., R. Gardner, H. Barik and K. Tunstall. 1963. Attitudinal and Cognitive Aspects of Intensive Study of Second Language. *Journal of Abnormal and Social Psychology*. 66. 358-368.
- Lamendella, J. 1977. General Principles of Neurofunctional Organization and their Manifestations in Primary and non-primary Acquisition. *Language Learning*. 27. 155-196.
- 1979. The Neurofunctional Basis of Pattern Practice. *TESOL Quarterly* 13. 5-19.
- Larsen-Freeman, D. 1975. The Acquisition of Grammatical Morphemes by Adult ESL Students. *TESOL Quarterly*. 9, 4. 409-419.
- 1976. An Explanation for the Morpheme Acquisition Order of Second Language Learners. *Language Learning*. 26. 125-134.
- 1980. *Discourse Analysis in Second Language Research*. Rowley, Mass.: Newbury House.
- 1983. Assessing Global Second Language Proficiency. In H. Seliger and M. Long (eds). 1983.
- Larsen-Freeman, D. and V. Strom. 1977. The Construction of a Second Language Acquisition Index of Development. *Language Learning*, vol. 27, no. 1. 123-134.
- 1976. An Explanation for the Morpheme Acquisition Order for Second Language Learners. *Language Learning*. 26. 125-134.
- Leech, G. 1971. *Meaning and the English Verb*. London: Longman.

- Lees, R. B. 1961. The Constituent Structure of Noun Phrases in *American Speech*. 35, 149-168.
- Lefebvre, C. 1975. Plural Agreement in Cuzco Quechua: Some aspects of Variation. Unpublished Ph.D. dissertation. University of California, Berkeley.
- Lenneberg, E. 1967. *Biological Foundations of Language*. New York: Wiley and Sons.
- Levinson, S. 1983. *Pragmatics*. London: Cambridge University Press.
- Levy, D. 1979. Communicative Goals and Strategies: Between Discourse and Syntax. In T. Givon ed.
- Lightbown, P. 1977. Consistency and Variation in the Acquisition of French: A study of First and Second Language Development. Unpublished Ph. D. thesis, Columbia University.
- 1985. Great Expectations: Second-Language Acquisition Research and Classroom Teaching. *Applied Linguistics*. 6. 173-89.
- Lindsey, J. 1975. Likelihood Analyses and Tests for Binary Data. *Applied Statistics*. 24. 1-16.
- Long, M. 1977. Teacher Feedback on Learners Error; Mapping Cognitions. In C. Brown, C. Yorio, and R Crymes (eds) *On TESOL '77*. Washington, D.C.: TESOL.
- 1980. Inside the "Black Box": Methodological Issues in Classroom Research on Language Learning. *Language Learning*. 30. 1-42.
- Long, M. H. and C. J. Sato. 1984. Methodological Issues in Interlanguage Studies: An Interactive Perspective. Paper presented at the seminar on Interlanguage in honour of S. P. Corder. MS Edinburgh University.
- Longacre, R. 1972. *Hierarchy and Universality of Discourse Constituents in New Guinea Languages: Discussion*. Washington, D.C. Georgetown University Press.
- 1976. *An Anatomy of Speech Notions*. Lisse: Peter de Ridder.
- Lukmani, Y. 1972. Motivation to Learn and Language Proficiency. *Language Learning*. 22. 261-273.
- Lyons, J. 1977. *Semantics*. Volumes I and II. London: Cambridge University Press.
- McDonough, S. 1981. *Psychology in Foreign Language Teaching*. London: Allen and Unwin.
- Magura, B. 1984. Style and Meaning in African English: A Sociolinguistic Analysis of South African and Zimbabwean English. Unpublished Ph. D. Thesis. University of Illinois at Urbana-Champaign.



- Maratsos, M. P. 1971. A Note on NPs made Definite by Entailment. *Linguistic Inquiry*. 1.
- 1973. Non-egocentric Communication Abilities in Preschool Children. *Child Development* 44. 697-700.
- 1976. *The Use of the Definite and Indefinite Reference in Young Children*. Cambridge University Press.
- Mason, C. 1971. The Relevance of Intensive Training in English as a Foreign Language for University Students. *Language Learning*. 21. 197-204.
- Mazrui, A. 1975. *The Political Sociology of the English Language*. The Hague: Mouton.
- Mazurkewich, I. 1982. Second Language Acquisition of the Dative Alternation and Markedness: The Best Theory. Ph.D. dissertation: Universite de Montreal.
- Mazurkewich, I. and L. White. 1984. The Acquisition of the Dative Alternation: Unlearning Overgeneralizations. *Cognition*. 16. 261-283.
- Mbaya, N. 1983. L'Anglais dans l'enseignement Zairois: 1956-1982. *Centre de Recherches Universitaires du Kivu, Collection Etudes*. 6. 1-55.
- Meisel, J., H. Clahsen and M. Pienemann. 1981. On Determining Developmental Stages in natural Second Language Acquisition. *Studies in Second Language Acquisition*. 3. 109-135.
- Mikes, M. 1967. Acquisition des Categories Grammaticales dans le langage de l'enfant. *Enfance*. 20 289-297.
- Morelong, M. and L. Hyman. 1977. Animacy, Objects and Clitics in Lesotho. *Studies in African Linguistics*. 8. 199-217.
- Naiman, N. 1974. The Use of Elicited Imitation in Second Language Acquisition Research. *Working Papers in Bilingualism*. 2. 1-34.
- Naro, A. and M. Lemle. Syntactic Diffusion. In S. Steever, C. Walker and S. Mufwene (eds). *Papers from the Parasession on Diachronic Syntax*. Chicago: Chicago Linguistic Society. 1976. 221-247.
- Naro, A. and M. Lemle. 1977. Syntactic Diffusion. *Ciencia e Cultura*. 29. 259-268.
- Neisser, U. 1967. *Cognitive Psychology*. New York: Appleton-Century-Crofts.
- Nehls, D. 1974. *Synchron-diachrone Untersuchungen zur Expanded Form im Englischen*. Muenchen: Huebner.
- Nelson, K. 1975. The Nominal Shift in Semantic-Syntactic Development. *Cognitive Psychology*. 7, 4. 461-479.
- Nemser, W. 1971. Approximative Systems of Foreign Language Learners. *I. R. A. L.* 9. 115-123.

- Neufeld, G. 1978. A Theoretical Perspective on the Nature of Linguistic Aptitude. *I. R. A. L.* 16. 15-26.
- 1979. Towards a Theory of Language Learning Ability. *Language Learning*. 29, 2. 227-241.
- Neuman, R. 1977. An Attempt to Define through Error Analysis the Intermediate ESL level at UCLA. Unpublished thesis. UCLA.
- Newmark, L. and D. Reibel. 1970. Necessity and Sufficiency in Language Learning. In M. Lester (ed). *Readings in Applied Transformational Grammar*. New York: Holt, Rinehart and Winston.
- Nie, H., C. Hull, J. Jenkins, K. Steinbrenner, D. Bent. 1975. *Statistical Package for Social Sciences - SPSS*. New York: McGraw Hill.
- Okanlawon, B. 1984. A Study of the Acquisition of Some aspects of English by Nigerian Learners. Unpublished Ph.D. thesis. University of Edinburgh.
- Oller, J. W. 1973. Cloze Tests of Second Language Proficiency and What They Measure. *Language Learning* 23, 1.
- 1978. How Important is Language Proficiency to I.Q. and Other Educational Tests?. In Oller, J. W. and Perkins, K. (eds) *Language in Education: Testing the Tests*. Rowley, Mass.: Newbury House.
- 1979. *Language Tests at School: A Pragmatic Approach*. London: Longman.
- 1983. (ed). *Issues in language Testing Research*. Rowley, Mass.: Newbury House.
- Oller, J. W. and K. Perkins. 1980 (eds). *Research in Language Testing*. Rowley, Mass.: Newbury House.
- Oller, J. W., K. Perkins and M. Murakami. 1980. Seven Types of Learner Variables in Relation to ESL Learning. In J. W. Oller and K. Perkins (eds). 1980.
- Oller J. W. and F. Hifonotis. 1980. Two Mutually Exclusive Hypotheses about Second Language Ability: Factor Analytic Studies of a Variety of Language Subtests. In J. W. Oller and K. Perkins (eds) 1980.
- Palmer, A. 1983. Compartmentalized and Integrated Control. In J. W. Oller (ed) 1983.
- Pavesi, M. 1986. A Study of Markedness in Second Language Acquisition. Unpublished Ph. D. thesis. University of Edinburgh.
- Perdue, C. (ed). 1982. *Second Language Acquisition by adult Immigrants. A field Manual*. Strasbourg, France: European Science Foundation.
- Pepinsky, H. 1974. A Metalanguage for Systematic Research on Human Communication via Natural Language. *Journal of the American Society for Information Science*. January-February 1974. 59-69.

- Perkins, K. and D. Larsen-Freeman. 1975. The Effect of Formal Language Instruction on the Order of Morpheme Acquisition. *Language Learning*. 25. 237-244.
- Perlmutter, D. M. 1968. Deep and Surface Structure Constraints in Syntax. Unpublished Ph.D. dissertation, MIT; published 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt Rinehart, and Winston.
- 1970. On the Article in English. In Bierwisch and Heidolph (eds).
- Peters, A. 1983. *The Units of Language Acquisition*. Cambridge University Press.
- Pica, T. 1983. Adult Acquisition of English as a Second Language under Different Conditions of Exposure. *Language Learning* 33. No. 4.
- 1984. Methods of Measuring Overall Morpheme Production Accuracy. Paper presented at the 18th annual TESOL Conference, Houston, Texas.
- Pindi, M. 1982. Some Common Errors committed in English by 2nd year Students of Economics at the University of Kinshasa. *RELAVI*. 5. 30-56. Published by the University of Kinshasa School of Languages.
- Platt, J., H. Weber and M. L. Ho. 1984. *The New Englishes*. London: Routledge and Kegan Paul.
- Platt, J. 1979. Variation and Implicational Relationships: Copula Realization in Singapore English. *General Linguistics*. 19, 1. 1-14.
- Politzer, R. and M. McGroarty. 1983. A Discrete point Test of Communicative Competence. *IRAL*. 21, 3. 179-191.
- Pollak, W. 1960. Studien zum Verbalsprekt im Franzosischen, Osterreichische ee der Wissenschaften, Bd. 233:5.
- Postal, P. 1966. On So-called Pronouns in English. *Georgetown Monograph Series* 19.
- Pride, J.ed.1982. *New Englishes*.Rowley, Mass.:Newbury House
- Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. 1972. *A Grammar of Contemporary English*. Longman.
- Radford, A. 1981. *Transformational Syntax*. Cambridge: Cambridge University Press.
- Ramirez, A. 1984. Pupil Characteristics and Performance on Linguistic and Communicative Language Measures. In C. Rivera (ed). *Communicative Competence Approaches to Assessment: Research and Application*. Multilingual Matters. 9: 82-106.
- Ravem, R. 1974. The Development of Wh-questions in First and Second Language Learners. In Richards (ed) 1974.
- Rea, P. 1986. Language Planning for Higher Education: Who Needs What and Why? In Centre of African Studies.1986.Language in Education in Africa.University of Edinburgh.

- Reinecke, J. 1969. Language and Dialect in Hawaii. In Tsuzaki ed. Reprint Honolulu: University of Hawaii Press.
- Richards, J. 1971. Error Analysis and Second Language Strategies. *Language Science*. 17. 12-22.
- 1973. *A Non-contrastive approach to Error Analysis. Focus on the Learner: Pragmatic Perspectives for the Language Teacher*. ed. by J. Oller and J. Richards. Rowley, Mass.: Newbury. 96-113.
- ed. 1978. *Understanding Second and Foreign Language Learning: Issues and Approaches*. Rowley, Mass.: Newbury House.
- Rivers, W. 1981. *Foreign Language Skills*. 2nd edition. Chicago: University of Chicago Press.
- Robson, C. 1973. *Experiment, Design and Statistics in Psychology*. Penguin.
- Rosansky, E. 1976. Methods and Morphemes in Second Language Acquisition Research. *Language Learning*. 26. 409-425.
- Rosenbaum, P. S. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, Mass.: MIT Press.
- Rousseau, P. and D. Sankoff. 1978. Advances in Variable Rule Methodology. In D. Sankoff (ed). 1978.
- Rutherford, W. 1982. Markedness in Second Language Acquisition Research. *Language Learning*. 32. 85-107.
- 1984 (ed). *Second Language Acquisition and Language Universals*. Amsterdam: John Benjamins.
- Rutherford, W. and M. Sharwood-Smith (forthcoming). *Readings in Pedagogical Grammar*. Rowley, Mass.: Newbury House.
- Ryan, E. 1983. Social Psychological Mechanisms Underlying Native Speaker Evaluations of Non-Native Speech. *Studies in Second Language Acquisition*. 5. 148-159.
- Sacks, H. 1972. An Initial Investigation of the Usability of Conversational Data for Doing Sociology. In D. Sudnow (ed). *Studies in Social Interaction*. New York: The Free Press.
- Sacks, H., E. Schegloff and G. Jefferson. 1974. A Simplest Systematics for the Analysis of Turn-taking in Conversations. *Language*. 50. 696-735.
- Salami, A. 1968. Defining a Standard Nigerian English. *Journal of the Nigerian English Studies Association*. 212. 99-106.
- Samarin, W. 1971. Salient and Substantive Pidginization. In D. Hymes ed.
- Sankoff, D. 1975. VARBRUL version 2. Unpublished Manuscript.

- 1978. ed. *Linguistic Variation: Models and Methods*. New York: Academic Press.
- Sankoff, D. and H. Cedergren. 1976. The Dimensionality of Grammatical Variation. *Language*. 52. no. 1. 163-178.
- Sankoff, D. and S. Laberge. 1978. Statistical Dependence among successive Occurrences of a Variable in Discourse. In D. Sankoff (ed) 1978.
- Sankoff, G. and P. Thibault. 1977. L'alternance entre les auxiliaries avoir et etre en francais parle a Montreal. *Langue Francaise* 34. 81-108.
- Sato, C. 1983. Task Variation in Interlanguage Phonology. In S. Gass and C. Madden (eds). *Proceedings of the 10th University of Michigan Conference on Applied Linguistics*. Rowley, Mass.: Newbury.
- Schachter, J. 1974. An Error in Error Analysis. *Language Learning*. 24. 205-214.
- Schegloff, E. 1979. The Relevance of Repair to Syntax for Conversation. In T. Givon ed.
- Schmidt, M. 1980. Coordinate Structures and Language Universals in Interlanguage. *Language Learning*. 30, 12. 397-416.
- Schumann, J. 1975. Affective Factors and the Problem of Age in Second Language Acquisition. *Language Learning*. 25. 209-236.
- 1976. Second Language Acquisition: The Pidginization Hypothesis. *Language Learning*. 26. 291-408.
- 1976. Social Distance as a Factor in Second Language Acquisition. *Language Learning*. 26. 135-144.
- 1978. The Relationship of Pidginization, Creolization and Decreolization to Second Language Acquisition. *Language Learning*. 28. 367- 380.
- 1982. Simplification, Transfer, and Reflexification as Aspects of Pidginization and Early Second Language Acquisition. *Language Learning*. 32. 337-392.
- Schumann, F. and J. Schumann. 1977. Diary of a language learner: an Introspective Study of Second Language Learning. In H. Brown, C. Yorio, and R. Crymes (eds) *On TESOL '77*. Washington, D.C.: TESOL.
- Searle, J. 1969. *Speech Acts: An Essay on the Philosophy of Language*. Cambridge: Cambridge University Press.
- 1979. *Expression and Meaning*. Cambridge: Cambridge University Press.
- Seliger, H. 1983. Learner Interaction in the Classroom and its effects on Language Acquisition. In H. Seliger and M. Long. (eds) 1983.
- Seliger, H. and M. Long. 1983. (eds) *Classroom Oriented Research*. Rowley, Mass.:

Newbury House.

- Selinker, L. 1969. Language Transfer. *General Linguistics*. 9, 2.
- 1972. Interlanguage. *J. R. A. L.* 10. 209-231.
- Selinker, L. and J. Lamendella. 1978. Two Perspectives on Fossilization in Interlanguage Learning. *Interlanguage Studies Bulletin*. 3. 143-191.
- Selinker, L., M. Swain and G. Dumas. 1975. The Interlanguage Hypothesis Extended to Children. *Language Learning*. 25. 139-152.
- Selinker, L., M. Trimble and L. Trimble. 1978. Rhetorical Function Shifts in EST. *TESOL Quarterly*. 12. No. 3.
- Sey, K. 1973. *Ghanaian English*. London: Macmillan.
- Sharwood-Smith, M. 1979. Strategies, Language Transfer and the Simulation of the Second Language Learner's Mental Operations. *Language Learning*. 29. 345-362.
- Simard, L., D. Taylor and H. Giles. 1976. Attribution Processes and Interpersonal Accommodation in a Bilingual Setting. *Language and Speech*. 19. 3374-387.
- Simukoko, Y. 1982. Some Aspects of the English of Bantu Speakers in Urban Primary Schools in Zambia. Unpublished Ph.D. thesis, University of Edinburgh.
- Smith, C. 1964. Determiners and Relative Clauses in a Generative Grammar of English. *Language* 40. 37-52.
- Smythe, P. R. Stennet and H. Feenstra. 1972. Attitude, Aptitude and Type of Instructional Program in Second Language Acquisition. *Canadian Journal of Behavioural Science*. 4. 307-321.
- Solan, L. 1978. Anaphora in Child Language. Unpublished Ph.D. thesis. University of Massachusetts.
- Sorace, A. 1985. Metalinguistic Knowledge and Language Use in Acquisition-poor Environments. *Applied Linguistics*. 6. No. 3.
- Sorenson, H. 1959. *Word Classes in English*. Copenhagen.
- Spencer, J.(ed).1971. *The English Language in Africa*. London:Longman.
- Spolsky, B. 1969. Attitudinal Aspects of Second Language Learning. *Language Learning*. 19, 3. 271-285.
- Stauble, A. M. and J. Schumann. 1983. Toward a Description of the Spanish-English Basilect, in Bailey, Long and Peck (eds) *Second Language Acquisition Studies*. Rowley, Mass.: Newbury House.
- Stauble, A. M. 1978. The Process of Decreolization: A Model for Second Language Development. *Language Learning*. 28. 29-54.

- 1981. A Comparison of a Spanish-English and Japanese-English L<sub>2</sub> Continuum: Verb Phrase Morphology. Paper presented at the European/ North American Workshop on Cross-Linguistic Second Language Acquisition Research, Lake Arrowhead, California.
- Stenson, N. 1975. Induced Errors. In J. Schumann and N. Stenson (eds). *New Frontiers in Second Language Learning*. Rowley, Mass.: Newbury House.
- Stern, H. 1967. *Foreign Language in Primary Education*. London: Oxford University Press.
- Stockwell, R., P. Schachter and B. Partee. 1973. *The Major Syntactic Structures of English*. New York: Holt, Rinehart and Winston.
- Swain, M. 1981. Time and Timing in Bilingual Education. *Language Learning*. 31. 1-15.
- Swain, M., G. Dumas and N. Naiman. 1974. Alternatives to Spontaneous Speech: Elicited Translation and Imitation as Indicators of Second Language Competence. *Working Papers in Bilingualism*. 3.
- Syngle, B. 1973. Second Language Acquisition Strategies of Children and Adults: A Cross-sectional Study. Ph.D. Dissertation. Louisiana State University.
- Tajfel, H. 1974. Social Identity and Intergroup Behaviour. *Social Science Information*. 13. 65-93.
- 1977. Differentiation between Social Groups. Quoted by H. Giles (ed) 1977. Also Available in H. Tajfel (ed) 1978. *Studies in the Social Psychology of Intergroup Relations*. London: Academic Press.
- Tarallo, F. and J. Myhill. 1983. Interference and Natural Language Processing in Second Language Acquisition. *Language Learning* 33. 55-76.
- Tarone, E. 1974. Speech Perception in Second Language Acquisition: A Suggested Model. *Language Learning* 24, 2. 223-233.
- 1980. Communication Strategies, Foreigner Talk and Repair in Interlanguage. *Language Learning* 30. 417-432.
- 1982. Systematicity and Attention in Interlanguage. *Language Learning*. 32. 69-84.
- 1983. On the Variability of Interlanguage Systems. *Applied Linguistics*. 4, 2. 143-163.
- Tarone, E., U. Frauenfelder and L. Selinker. 1976. Systematicity / Variability and Stability / Instability in Interlanguage Systems. In D. Brown (ed). *Papers in Second Language Acquisition*. Ann Arbor: Language Learning.
- Taylor, B. 1975. The Use of Overgeneralization and Transfer Learning Strategies by Elementary and Intermediate University Students learning ESL. In

M. Burt and H. Dulay (eds) *New Directions in Second Language Learning, Teaching and Bilingual Education*. Washington, D.C.: TESOL.

- Taylor, W. L. 1953. Cloze Procedure: A New Tool for Measuring Readability. *Journalism Quarterly*. 33.
- Thompson, S. 1978. Modern English from a Typological Point of View: Some Implications of the Function of Word Order. *Linguistische Berichte*. 54. 19-35.
- Thrane, T. 1980. *Referential-Semantic Analysis: Aspects of a Theory of Linguistic Reference*. Cambridge: Cambridge University Press.
- Torgerson, W. 1958. *Theory and Methods of Scaling*. New York: Wiley.
- Traugott, E. and M. Pratt. 1980. *Linguistics for Students of Literature*. New York: Harcourt Brace.
- Upshur, J. 1962. Language Proficiency Testing and the Contrastive Analysis Dilemma. *Language Learning*. 12, 2.
- 1968. Four Experiments on the Relation between Foreign Language Teaching and Learning. *Language Learning*. 18. 111-124.
- Upshur, J. and T. Homberg. 1983. Some Relations among Language tests at successive ability levels. In *Issues in Language Testing Research*. Rowley, Mass.: Newbury House.
- Van Mettre, P. 1972. Syntactic Characteristics of Selected Bilingual Children. Ph. D. dissertation. University of Arizona.
- Varadi, T. 1983. Strategies of Target Language Learner Communication: Message Adjustment. In Faerch and Kasper eds.
- Wagner-Gough, J. 1975. Comparative Studies in Second Language Learning. M.A. Thesis. UCLA.
- Wagner-Gough, J. and E. Hatch. 1975. The Importance of Input Data in Second Language Acquisition Studies. *Language Learning*. 25. 297-308.
- Wallace, C. 1986. *Learning to Read in a Multicultural Society*. Oxford: Pergamon.
- Weissenburger, J. 1976. A Choice of Words: Two-year-old speech from a Situational point of view. *Journal of Child Language*. 3. 275-281.
- Whinnom, K. 1971. Linguistic Hybridization and the Special Case of Pidgins and Creoles. In D. Hymes ed.
- White, L. 1987. Against Comprehensible Input: The input hypothesis and the development of Second-language Competence. *Applied Linguistics*. 8, 2. 95-109.
- Widdowson, H. 1979. Rules and Procedures in Discourse Analysis. In T. Myers (ed) *The Development of Conversation and Discourse*. Edinburgh:



Edinburgh University Press.

- Wilkins, D. 1968. Review of A. Valdman (ed) *Trends in Language Teaching*. *IRAL*. VI, 1.
- Winer, B. 1971. *Statistical Principles in Experimental Design*. 2nd Edition. New York: McGraw-Hill.
- Wittgenstein, L. 1958. *Philosophical Investigations*. Oxford: Blackwell.
- Wode, H. 1976. Developmental Sequences in Naturalistic Language Teaching. *Working Papers on Bilingualism*. 11.
- Wolfe, P. 1967. The Operation of Pronominalization within the NP, with particular reference to English. Unpublished paper, U.C.L.A. (Pegs paper £ 52).
- 1968. Definite and Indefinite Pronominalization in English. LSA Summer meeting.
- Zobl, H. 1980. The Formal and Developmental Selectivity of L<sub>1</sub> Influence on L<sub>2</sub> Acquisition. *Language Learning* 30. 43-58.
- 1983a. Markedness and the Projection Problem. *Language Learning*. 33. 293-313.
- 1983b. Grammars in search of Input and Intake. Paper presented at the 10th University of Michigan Conference on Applied Linguistics.
- 1984a. Aspects of Reference and the Pronominal Syntax Preference in the Speech of Young Child L<sub>2</sub> Learners. In R. Andersen (ed). 1984.
- 1984b. Cross-language Generalisations and the Contrastive Dimension of the Interlanguage Hypothesis. In A. Davies et al (eds). 1984.
- Zuengler, J. 1982. Kenyan English. In B. B. Kachru (ed) 1982.
- Zughoul, M. and O. Kambal. 1983. Objective Evaluation and EFL Composition. *I. R. A. L.* 21, 2. 87-101.
- Zydattis, W. 1974. Tense and Aspect in the English of German-speaking Learning. Ph.D. Dissertation. University of Edinburgh.
- 1976. Learning Problem Expanded Form - A Performance Analysis. *I. R. A. L.* 14, 4. 351-371.

Table A

Implicational scale for the Pupils' Use of  
Definite Reference *THE* in three environments

S.No.	THE + PREP	THE + AGT	THE + OBJ	
162	+	+	+	
147	+	+	+	
165	+	+	+	
141	+	+	+	
144	+	+	+	
160	+	+	+	
148	+	+	+	
159	+	+	+	
116	+	+	-	
135	+	+	-	
137	+	+	-	
112	+	+	-	
142	+	+	-	
109	+	+	-	
114	+	+	-	
151	+	+	-	
131	+	+	-	
157	-	+	⊕	
145	-	+	⊕	
127	+	-	⊕	
139	+	-	⊕	
166	-	+	-	
130	-	+	-	
110	+	-	-	
155	+	-	-	
111	+	-	-	
126	+	-	-	
161	+	-	-	
125	+	-	-	
103	+	-	-	
149	+	-	-	
163	+	-	-	
134	+	-	-	
152	+	-	-	
106	+	-	-	
136	-	-	⊕	
115	-	-	⊕	
118	-	-	⊕	
133	-	-	⊕	
158	-	-	-	
138	-	-	-	
Correct	32	21	16	= 69
Errors	8	2	8	= 18

Coefficient Rep. = 0.91  
Coefficient Scal. = 0.86

Table B

**Implicational Scale for the Undergraduates' Use of  
Definite Reference *THE* in three environments**

S.No.	THE + PREP	THE + AGT	THE + OBJ	
96	+	+	+	
87	+	+	+	
91	+	+	+	
89	+	+	+	
85	+	+	+	
84	+	+	-	
86	+	+	-	
94	-	+	⊕	
79	+	-	⊕	
80	+	-	-	
90	+	-	-	
83	+	-	-	
95	-	-	+	
78	-	-	-	
97	-	-	-	
88	-	-	-	
92	-	-	-	
81	-	-	-	
93	-	-	-	
82	-	-	-	
Correct	11	8	8	= 27
Errors	2	1	3	= 6

Coefficient rep. = 0.90

Coefficient scal. = 0.82

Table C

**Implicational Scale for the T-Teachers' Use of  
Definite Reference *THE* in three environments**

S.No.	THE + PREP	THE + AGT	THE + OBJ		
35	+	+	+		
71	+	+	+		
7	+	+	+		
67	+	+	+		
39	+	+	+		
54	+	+	+		
74	+	+	+		
46	+	+	+		
55	+	+	+		
47	+	+	+		
29	+	+	+		
48	+	+	+		
42	+	+	-		
42	+	+	-		
41	+	+	-		
15	+	+	-		
66	+	+	-		
68	+	+	-		
38	+	+	-		
65	+	+	-		
49	+	+	-		
62	+	+	-		
6	+	+	-		
69	+	+	⊕		
72	-	+	⊕		
5	+	-	⊕		
58	+	-	⊕		
40	+	-	⊕		
24	+	-	⊕		
53	+	-	⊕		
43	-	+	-		
23	-	+	-		
37	-	+	-		
75	+	-	-		
76	+	-	-		
44	+	-	-		
70	+	-	-		
45	+	-	-		
51	+	-	-		
26	+	-	-		
64	+	-	-		
2	+	-	-		
57	-	-	⊕		
Correct	37	27	20	=	84
Errors	5	5	8	=	18

Coefficient rep. = 0.92

Coefficient scal. = 0.87

Table D

**Implicational Scale for the Pupils' Use of  
Indefinite Reference *A*, *AN* in three environments**

S. No.	AN + AGT	AN + OBJ	AN + PREP
135	+	+	+
136	+	+	+
149	+	+	+
157	+	+	+
160	+	+	+
106	+	+	+
119	+	+	+
116	+	+	+
134	+	+	-
105	+	+	-
141	+	+	-
142	+	+	-
147	+	+	-
155	+	+	-
124	+	+	-
161	+	+	-
165	+	+	-
118	+	+	-
120	+	+	-
108	+	+	-
128	+	+	-
130	+	+	-
121	+	+	-
158	-	+	⊕
167	-	+	⊕
143	-	-	⊕
144	-	-	⊕
166	-	-	⊕
131	-	-	⊕
103	-	+	-
156	-	-	-
137	+	-	-
139	+	-	-
140	+	-	-
145	+	-	-
162	+	-	-
150	+	-	-
151	+	-	-
152	+	-	-
122	+	-	-
159	+	-	-
127	+	-	-
163	+	-	-
164	+	-	-
102	+	-	-
104	+	-	-
107	+	-	-
109	+	-	-
110	+	-	-
111	+	-	-
114	+	-	-
123	+	-	-
101	+	-	-
126	+	-	-
129	+	-	-
125	+	-	-
132	+	-	-
133	+	-	-
138	-	-	-
146	-	-	-
112	-	-	-
113	-	-	-
115	-	-	-
117	-	-	-
Correct	56	28	8 = 92
Errors	4	4	6 = 14
Coefficient rep. = 0.92			
Coefficient scal = 0.87			

Table E

Implicational Scale for the Undergraduates' Use of Indefinite Reference *A*, *AN* in three environments

S.No.	AN + AGT	AN + OBJ	AN + PREP	
77	+	+	+	
80	+	+	+	
96	+	+	-	
85	+	+	-	
91	+	+	-	
94	+	+	-	
95	+	+	-	
83	-	+	-	
97	+	-	⊕	
84	+	-	⊕	
93	+	-	⊕	
89	+	-	⊕	
81	+	-	⊕	
87	+	-	-	
88	+	-	-	
82	+	-	-	
79	+	-	-	
86	-	-	-	
90	-	-	-	
78	-	-	-	
92	-	-	-	
Correct	16	8	7	= 31
Errors	1	0	5	= 6

Coefficient rep. = 0.90  
 Coefficient scal = 0.80

Table F

**Implicational Scale for the T-Teachers' Use of Indefinite Reference *A*, *AN* in three environments**

S. No.	AN + AGT	AN + OBJ	AN + PREP
74	+	+	+
41	+	+	+
42	+	+	+
46	+	+	+
48	+	+	+
24	+	+	+
61	+	+	+
62	+	+	+
2	+	+	+
6	+	+	+
20	+	+	+
69	+	+	-
76	+	+	-
71	+	+	-
8	+	+	-
66	+	+	-
67	+	+	-
43	+	+	-
54	+	+	-
55	+	+	-
64	+	+	-
7	+	+	-
9	+	+	-
26	+	+	-
27	+	+	-
15	+	+	-
32	+	+	-
22	+	+	-
72	-	+	⊕
75	+	-	⊕
70	+	-	⊕
68	+	-	⊕
37	+	-	⊕
45	+	-	⊕
16	+	-	⊕
18	+	-	⊕
58	+	-	⊕
1	+	-	⊕
29	+	-	⊕
30	+	-	⊕
10	-	+	-
73	+	-	-
65	+	-	-
38	+	-	-
39	+	-	-
40	+	-	-
44	+	-	-
47	+	-	-
50	+	-	-
52	+	-	-
53	+	-	-
56	+	-	-
59	+	-	-
60	+	-	-
63	+	-	-
33	+	-	-
35	+	-	-
5	+	-	-
23	+	-	-
25	+	-	-
13	+	-	-
31	+	-	-
17	+	-	-
19	+	-	-
21	+	-	-
28	+	-	-
57	-	-	-
11	-	-	-
12	-	-	-
3	-	-	-
14	-	-	-
49	-	-	-
51	-	-	-
34	-	-	-
4	-	-	-
36	-	-	-
Correct	64	30	23 = 117
Errors	2	11	12 = 25

Coefficient rep. = 0. 89  
Coefficient scal. = 0. 78

Table G

**Implicational Scale for the Pupils' Use of Definite  
Reference THE<sub>2</sub> in three environments**

S.No.	THE2+AGT	THE2+PREP	THE2+OBJ	
165	+	+	+	
162	+	+	+	
112	+	+	-	
144	+	+	-	
164	+	+	-	
147	+	+	-	
155	+	+	-	
127	+	+	-	
148	+	-	⊕	
111	+	-	⊕	
141	-	+	-	
135	-	+	-	
108	-	+	-	
133	-	+	-	
160	-	+	-	
137	-	+	-	
110	-	+	-	
124	-	+	-	
109	-	+	-	
134	+	-	-	
131	+	-	-	
166	+	-	-	
103	+	-	-	
119	+	-	-	
139	+	-	-	
156	+	-	-	
149	+	-	-	
116	+	-	-	
126	+	-	-	
161	+	-	-	
159	-	-	⊕	
157	-	-	⊕	
142	-	-	-	
136	-	-	-	
130	-	-	-	
140	-	-	-	
132	-	-	-	
114	-	-	-	
163	-	-	-	
106	-	-	-	
120	-	-	-	
138	-	-	-	
167	-	-	-	
151	-	-	-	
105	-	-	-	
101	-	-	-	
Correct	21	17	6	= 44
Errors	11	2	4	= 17

Coefficient Rep. = 0.91 Coefficient Scal. = 0.88



n.b. There is no scale for the three items put together, thus no statistics are available for this scale.

Table H

**Implicational Scale for the Undergraduates' Use of  
Definite Reference THE<sub>2</sub> in three environments**

S.No.	THE2+AGT	THE2+PREP	THE2+OBJ
94	+	+	-
95	+	+	-
91	-	+	-
79	-	+	-
96	+	-	-
80	+	-	-
86	+	-	-
78	-	-	⊕
93	-	-	⊕
8	-	-	-
86	-	-	-
89	-	-	-
85	-	-	-
90	-	-	-
84	-	-	-
87	-	-	-
77	-	-	-
92	-	-	-
88	-	-	-
97	-	-	-
82	-	-	-

Table I

Implicational scale for the T-Teachers' Use of  
Definite Reference THE<sub>2</sub> in three environments

S.No.	THE2+AGT	THE2+PREP	THE2+OBJ	
5	+	+	+	
7	+	+	-	
67	+	+	-	
47	+	+	-	
26	+	+	-	
15	+	-	⊕	
6	+	-	⊕	
54	-	+	-	
2	-	+	-	
74	-	+	-	
66	-	+	-	
48	+	-	-	
57	+	-	-	
44	+	-	-	
69	+	-	-	
71	+	-	-	
40	+	-	-	
41	+	-	-	
49	+	-	-	
51	+	-	-	
35	+	-	-	
53	+	-	-	
30	+	-	-	
62	+	-	-	
20	-	-	⊕	
11	-	-	⊕	
1	-	-	⊕	
43	-	-	-	
39	-	-	-	
42	-	-	-	
45	-	-	-	
75	-	-	-	
68	-	-	-	
76	-	-	-	
46	-	-	-	
23	-	-	-	
65	-	-	-	
70	-	-	-	
58	-	-	-	
24	-	-	-	
Correct	20	9	6	= 35
Errors	7	2	5	= 14

Coefficient rep. = 0.94

Coefficient scal. = 0.93

Table J

**Developmental Stages Among 3 Articles Within 3  
Environments: Multidimensional Implicational Patterns.**

S.No.	A, AN			THE			THE2		
	AGT	OBJ	PREP	AGT	OBJ	PREP	AGT	OBJ	PREP
STAGE V									
165	1.00+	0.90+	0.40-	1.00+	0.80+	1.00+	1.00+	1.00+	0.86+
74	1.00+	1.00+	1.00+	1.00+	0.80+	0.80+	0.25-	0.33-	1.00+
160	1.00+	1.00+	0.80+	0.91+	0.80+	1.00+	0.50-	0.00-	0.86+
STAGE IV									
157	1.00+	1.00+	0.80+	0.82+	0.80+	0.50-	0.75-	1.00+	0.71-
6	1.00+	0.80+	1.00+	1.00+	0.06+	0.80+	1.00+	1.00+	0.71-
41	1.00+	0.90+	1.00+	0.91+	0.40+	1.00+	1.00+	0.00-	0.71-
135	1.00+	0.90+	0.80+	0.90+	0.60+	1.00+	0.75-	0.00-	0.86+
7	1.00+	0.90+	0.60-	0.91+	1.00+	1.00+	1.00+	0.67-	0.86+
147	1.00+	0.90+	0.40-	0.82+	0.80+	1.00+	1.00+	0.67-	1.00+
91	1.00+	0.80+	0.40-	0.91+	1.00+	1.00+	1.00+	0.00-	0.71-
96	1.00+	0.90+	0.60-	0.91+	0.80+	1.00+	0.75-	0.33-	1.00+
141	1.00+	0.90+	0.60-	0.90+	0.80+	1.00+	0.75-	0.00-	0.86+
2	1.00+	0.90+	0.80+	0.73-	0.60-	0.80+	0.75-	0.67-	0.86+
80	1.00+	0.80+	1.00+	0.55-	0.40-	0.80+	0.50-	0.67-	0.86+
66	1.00+	0.90+	0.40-	0.91+	0.60-	1.00+	0.75-	0.00-	0.86+
95	1.00+	0.90+	0.40-	0.73-	1.00-	0.50+	1.00+	0.67-	0.86+
155	1.00+	1.00+	0.40-	0.73-	0.60-	1.00+	1.00+	0.33-	0.86+
144	1.00+	0.56-	0.80+	0.82+	0.80+	1.00+	1.00+	0.33-	0.86+
162	1.00+	0.67-	0.60-	1.00+	1.00+	1.00+	1.00+	1.00+	0.86+
5	1.00+	0.67-	0.60-	0.73-	0.80+	1.00+	1.00+	1.00+	0.86+
STAGE III									
95	1.00+	0.90+	0.80+	0.27-	0.40-	0.50-	1.00+	0.33-	0.57-
77	1.00+	0.80+	0.80+	0.73-	0.40-	0.50-	0.50-	0.00-	0.43-
118	1.00+	0.80+	0.20-	0.73-	0.80+	0.50-	0.25-	0.33-	0.71-
43	1.00+	0.80+	0.60-	0.80+	0.60-	0.50-	0.75-	0.00-	0.40-
142	1.00+	1.00+	0.60-	0.91+	0.60-	1.00+	0.75-	0.00-	0.57-
134	1.00+	0.80+	0.40-	0.55-	0.40-	0.75+	1.00+	0.00-	0.43-
93	1.00+	0.67-	0.80+	0.27-	0.20-	0.00-	0.50-	1.00+	0.29-
45	1.00+	0.67-	0.80+	0.55-	0.20-	0.75+	0.75-	0.33-	0.71-
18	1.00+	0.44-	0.80+	0.55-	0.80+	0.00-	0.50-	0.33-	0.29-
37	1.00+	0.44-	0.80+	0.82+	0.40-	0.25-	0.75-	0.67-	0.57-
159	1.00+	0.44-	0.20-	0.82+	0.80+	0.75+	0.75-	1.00+	0.57-
87	1.00+	0.44-	0.40-	0.82+	0.80+	1.00+	0.50-	0.33-	0.71-
68	1.00+	0.67-	0.80+	0.91+	0.60-	1.00+	0.75-	0.67-	0.57-
79	1.00+	0.67-	0.60-	0.73-	0.80+	1.00+	1.00+	0.67-	0.57-
139	1.00+	0.33-	0.20-	0.45-	0.80+	0.75+	1.00+	0.33-	0.71-
145	1.00+	0.56-	0.00-	0.91+	0.80+	0.25-	0.50-	0.00-	0.43-
114	1.00+	0.56-	0.60-	0.91+	0.60-	0.75+	0.75+	0.67-	0.57-
133	1.00+	0.44-	0.60-	0.18-	0.80+	0.50-	0.50-	0.00-	0.86+
126	1.00+	0.56-	0.60-	0.36-	0.60-	1.00+	1.00+	0.00-	0.71-
72	0.00-	0.80+	0.80+	0.91+	0.80+	0.50-	0.50-	0.33-	0.71-
112	0.00-	0.56-	0.20-	0.82+	0.60-	1.00+	1.00+	0.33-	0.86+
103	0.00-	0.89+	0.00-	0.73-	0.60-	0.75+	1.00+	0.00-	0.57-

TABLE J CONTINUED

S.No.	A, AN			THE			THE2		
	AGT	OBJ	PREP	AGT	OBJ	PREP	AGT	OBJ	PREP
STAGE II									
8	1.00+	0.90+	0.40-	0.73-	0.20-	0.25-	0.00-	0.33-	0.43-
22	1.00+	0.80+	0.60-	0.00-	0.00-	0.25-	0.25-	0.67-	0.71-
9	1.00+	0.80+	0.60-	0.36-	0.20-	0.25-	0.50-	0.33-	0.29-
32	1.00+	0.80+	0.40-	0.73-	0.60-	0.50-	0.50-	0.33-	0.71-
27	1.00+	0.80+	0.40-	0.64-	0.40-	0.50-	0.50-	0.67-	0.71-
16	1.00+	0.00-	0.80+	0.00-	0.00-	0.00-	0.25-	0.00-	0.43-
97	1.00+	0.56-	1.00+	0.64-	0.60-	0.50-	0.25-	0.67-	0.57-
163	1.00+	0.56-	0.20-	0.64-	0.60-	0.80+	0.75-	0.67-	0.57-
125	1.00+	0.33-	0.40-	0.36-	0.60-	1.00+	0.25-	0.67-	0.71-
123	1.00+	0.67-	0.40-	0.73-	0.40-	0.50-	0.50-	0.00-	0.71-
140	1.00+	0.67-	0.20-	0.55-	0.40-	0.50-	0.75-	0.67-	0.71-
33	1.00+	0.56-	0.40-	0.36-	0.40-	0.25-	0.50-	0.67-	0.43-
60	1.00+	0.67-	0.40-	0.64-	0.40-	0.50-	0.25-	0.00-	0.43-
21	1.00+	0.56-	0.40-	0.45-	0.60-	0.25-	0.25-	0.00-	0.57-
13	1.00+	0.10-	0.60-	0.55-	0.60-	0.25-	0.25-	0.00-	0.70-
153	1.00+	0.22-	0.20-	0.45-	0.40-	0.50-	0.25-	0.67-	0.71-
113	1.00+	0.00-	0.60-	0.18-	0.60-	0.00-	0.25-	0.00-	0.43-
129	1.00+	0.56-	0.40-	0.64-	0.40-	0.25-	0.00-	0.00-	0.43-
150	1.00+	0.67-	0.40-	0.73-	0.20-	0.25-	0.50-	0.00-	0.29-
158	0.00-	1.00+	0.80+	0.36-	0.00-	0.50-	0.50-	0.33-	0.43-
167	0.00-	0.80+	0.80+	0.73-	0.60-	0.50-	0.50-	0.67-	0.71-
STAGE I									
11	0.00-	0.10-	0.60-	0.45-	0.60-	0.25-	0.25-	1.00+	0.43-
90	0.00-	0.67-	0.40-	0.45-	0.00-	0.80+	0.50-	0.67-	0.57-
115	0.00-	0.67-	0.00-	0.55-	0.80+	0.50-	0.50-	0.00-	0.57-
10	0.00-	0.80+	0.60-	0.64-	0.40-	0.50-	0.50-	0.33-	0.43-
92	0.00-	0.67-	0.20-	0.27-	0.20-	0.25-	0.50-	0.00-	0.43-

Table K

Scheffe Tests a 2-way Interaction Between Verb Types and  
The Semantic Features of the Main Clause Subject

	1 Promise x Inan. Subject	2 Promise x Animate Subject	3 Promise x Human Subject	4 Tell x Animate Subject	5 Ask x Human Subject	6 Tell Inan. Subject	7 Ask Animate Subject	8 Tell Human Subject	9 Ask Inan. Subject
mean =	246.95	256.24	268.95	316.86	324.21	337.90	353.06	405.87	425.92
1. Promise x Inan. S	--			** 69.91	** 77.26	** 90.95	** 106.11	** 158.92	** 178.97
2. Promise x An. S		--		** 60.62	** 67.97	** 81.66	** 96.82	** 149.63	** 169.68
3. Promise x Human S			--		* 55.26	** 68.95	** 84.11	** 139.92	** 156.97
4. Tell x An. S				--				** 89.01	** 109.09
5. Ask x Human S					--			** 81.66	** 101.71
6. Tell x Inan. S						--		** 67.97	** 88.02
7. Ask x An. S							--		** 72.86

\*\* p &lt; .01

\* p &lt; .05

Table L

Scheffe Tests for a 2-way Interaction between Verb Types  
and the Semantic Features of the Main Clause Object

	1 Promise Animate Object	2 Promise Human Object	3 Promise Inan. Object	4 Ask Inan. Object	5 Tell Inan. Object	6 Tell Animate Object	7 Tell Human Object	8 Ask Human Object	9 Ask Animate Object
mean =	221.03	246.95	304.16	329.10	356.97	370.17	378.98	384.84	389.24
1. Promise x An. Object	--		** 83.13	** 108.07	** 135.94	** 149.14	** 157.95	** 163.81	** 168.21
2. Promise x Human Object		--	** 57.21	** 82.15	** 110.02	** 123.22	** 132.03	** 137.89	** 142.29
3. Promise x Inan. Object			--		** 52.81	** 66.01	** 74.82	** 80.68	** 85.08
4. Ask x Inan. Object				--			** 49.88	** 55.74	** 60.14

\*\* P &lt; .01







Table O

**Implicational Scales for Individuals [Human-animate  
-inanimate]**

66 Pupils

S No.	HIN	AIN	IHU	HHU	HAN	IIN	IAN	AHU	AAN
133	+	+	+	+	+	+	+	+	+
150	+	+	+	+	+	+	+	+	+
128	+	+	+	+	+	+	+	+	+
141	+	+	+	+	+	+	+	+	+
142	+	+	+	+	+	+	+	+	+
160	+	+	+	+	+	+	+	+	+
108	+	+	+	+	+	+	+	-	+
118	+	+	+	+	+	+	-	+	+
134	+	+	+	+	+	-	-	-	-
151	+	+	+	+	-	-	-	-	-
144	+	+	+	-	+	+	+	+	-
130	+	+	+	-	-	+	+	-	+
131	+	+	-	+	+	-	+	-	-
165	+	+	-	+	-	+	+	-	-
155	+	+	-	-	-	+	-	-	-
143	-	+	+	+	+	+	+	+	+
120	-	+	+	+	+	+	+	+	+
159	-	+	+	-	+	+	-	-	+
116	-	+	+	-	-	-	-	+	+
111	-	+	+	-	-	-	-	-	-
161	-	+	+	-	-	-	-	-	-
153	+	+	-	-	+	+	+	+	-
167	-	+	-	+	+	+	+	+	+
103	-	+	-	-	+	+	-	-	-
158	-	+	-	-	-	-	+	-	-
123	-	+	-	-	-	-	-	-	-
102	-	+	-	-	-	-	-	-	-
156	-	+	-	-	-	-	-	-	-
145	+	-	+	+	+	+	+	+	+
163	+	-	+	+	-	+	-	-	+
147	+	-	+	+	-	-	-	-	-
115	+	-	+	-	-	-	-	-	-
129	+	-	+	-	-	-	-	-	-
127	+	-	+	-	+	+	+	+	+
148	+	-	-	+	-	+	+	-	-
140	+	-	-	-	+	-	-	+	+
126	+	-	-	-	-	-	+	-	+
109	+	-	-	-	-	-	-	-	+
101	+	-	-	-	-	-	-	-	-
137	-	-	+	+	-	-	-	-	-
Cor.	28	38	25	24	22	22	21	17	17 = 204
Err.	21	0	10	15	14	14	14	11	12 = 111

Coef. rep. = 0.81

Coef. scal. = 0.71

Table P

Implicational Scales for Individuals [Human-animate  
-inanimate]

21 Undergrads

S No.	AAN	AHU	IHU	HHU	IAN	HIN	AIN	IIN	HAN
89	+	+	+	+	+	+	-	-	-
85	+	+	+	+	+	+	+	-	+
77	+	+	+	+	-	+	-	-	-
94	-	+	+	+	-	+	+	-	-
96	+	+	+	-	+	+	-	+	-
79	+	+	-	+	-	-	-	+	-
91	+	+	-	-	+	+	+	-	+
80	+	+	-	-	+	-	-	-	-
97	+	+	-	-	-	-	-	+	-
88	-	+	-	-	-	-	-	-	+
87	+	-	+	+	+	-	+	+	-
82	+	-	+	-	-	-	-	-	-
83	+	-	+	-	+	+	+	-	-
81	+	-	-	-	-	-	+	+	-
78	+	-	-	-	-	-	-	-	-
93	-	-	+	+	-	-	-	-	-
95	-	-	-	+	-	-	-	+	-
84	-	-	-	-	+	-	+	-	-
92	-	-	-	-	-	-	-	-	+
86	-	-	-	-	-	-	-	-	-
90	-	-	-	-	-	-	-	-	-
Cor.	13	10	9	8	8	7	7	6	4 = 72
Err.	6	0	5	5	7	4	7	6	4 = 44

Coef. rep. = 0.77  
 Coef. scal. = 0.63

Table Q

Implicational Scales for Individuals [Animacy Hierarchy]

76 Postgrads

S No.	AAN	IIN	HHU	IAN	HIN	AHU	AIN	IHU	HAN
66	+	+	+	+	+	+	+	+	+
56	+	+	+	+	+	+	+	+	+
49	+	+	+	+	+	+	+	+	-
54	+	+	+	+	+	+	+	-	-
50	+	+	+	+	+	-	+	+	-
10	+	+	+	+	-	+	+	+	-
75	+	+	+	+	-	-	+	+	-
32	+	+	-	+	+	-	+	+	-
2	+	+	-	+	+	-	+	+	-
31	+	+	+	-	-	+	-	-	-
59	+	+	-	+	+	-	+	+	-
29	+	+	-	+	+	-	-	+	-
35	+	+	-	-	-	-	-	-	+
41	+	+	-	-	-	-	-	+	-
55	-	+	+	+	+	+	+	+	-
71	-	+	+	+	+	-	-	-	-
61	-	+	+	+	-	-	-	-	-
73	-	+	+	-	-	-	-	-	-
23	-	+	+	-	-	-	-	-	-
38	-	+	+	-	-	-	-	-	-
6	-	+	-	+	+	-	-	-	-
58	-	+	-	-	-	-	-	-	+
19	-	+	-	-	-	-	-	-	-
64	-	+	-	-	-	-	-	-	-
3	-	+	-	-	-	-	-	-	-
68	+	-	+	+	+	+	+	+	+
74	+	-	+	+	-	-	-	-	-
70	+	-	+	-	+	+	+	+	+
67	+	-	+	+	+	-	+	-	+
15	+	-	+	-	+	-	+	-	+
52	+	-	+	+	-	-	+	+	-
27	+	-	-	+	-	-	+	-	-
69	+	-	-	-	-	+	+	+	-
7	+	-	-	-	-	+	-	-	+
60	+	-	-	-	-	+	-	-	+
18	+	-	-	-	-	-	+	-	-
53	+	-	-	-	-	-	-	-	-
17	+	-	-	-	-	-	-	-	-
42	-	-	+	-	-	-	-	-	-
16	-	-	+	-	-	-	-	-	-
Cor.	27	25	23	23	22	20	19	19	11 = 189
Err.	26	0	17	15	18	16	15	16	9 = 132

Coef. Rep. = 0.81

Coef. Scal. = 0.74

Table R

## Results of TELL Sentences: NP1 and NP2 have equal status

SENTENCE TYPE	Level1 High school		Level2 Undergrad.		Level3 Teacher Tr.		Total	
	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.
1.) Human NP x Human  e.g. The boy told the girl to leave	17 (26%)	49 (74%)	9 (43%)	12 (57%)	26 (34%)	50 (66%)	52 (32%)	111 (68%)
2.) Animate x Animate  e.g. The dog told the lion to wait.	19 (29%)	47 (71%)	2 (10%)	19 (57%)	18 (24%)	58 (76%)	39 (24%)	124 (76%)
3.) Inan. x Inan.  e.g. The pen told the book to stop.	12 (18%)	54 (82%)	12 (57%)	9 (43%)	17 (22%)	59 (78%)	41 (25%)	122 (75%)
TOTAL/LEVEL	48 (24%)	150 (76%)	23 (37%)	40 (63%)	61 (27%)	167 (73%)	132 (27%)	357 (73%)
$\chi^2 = 13.465$ ; DF = 6; *P < .036								

- % in brackets indicates row percentage of errors vs. correct response per sentence type

Table S

Results of TELL Sentences: NP2 is higher on Topicality  
Hierarchy [convergence]

SENTENCE TYPE	Level1 High school		Level2 Undergrad.		Level3 Teacher Tr.		Total	
	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.
1.) Inanimate x Animate  e.g. The house told the chicken to go away.	5 - ( 8%	61 92%)	2 - (10%	19 90%)	7 - ( 9%	69 91%)	14 - ( 9%	149 91%)
2.) Inanimate x Human  e.g. The car told the soldier to stop.	5 - ( 8%	61 92%)	0 - (0%	21 100%)	3 - ( 4%	73 96%)	8 - ( 5%	155 95%)
3.) Animate x Human  e.g. The chicken told the child to have a biscuit.	23 - (35%	43 65%)	3 - (14%	18 86%)	17 - (22%	59 78%)	43 - (26%	120 74%)
TOTAL/LEVEL	33 - (17%	165 83%)	5 - ( 8%	58 92%)	27 - (12%	201 88%)	65 - (13%	424 87%)
$\chi^2 = 6.0; DF = 6; *P < .036$								

- % in brackets indicates row percentage of errors vs. correct response per sentence type

Table T

Results of TELL Sentences: NP1 is higher on the Topicality Hierarchy [conflict]

SENTENCE TYPE	Level1 High school		Level2 Undergrad.		Level3 Teacher Tr.		Total	
	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.
1.) Animate x Inanimate e.g. The chicken told the house to go away.	18 (27%)	48 (73%)	5 (24%)	16 (76%)	30 (39%)	46 (61%)	53 (33%)	110 (67%)
2.) Human x Inanimate e.g. The soldier told the car to stop.	24 (36%)	42 (64%)	10 (4%)	11 (52%)	31 (41%)	45 (59%)	65 (40%)	98 (60%)
3.) Human x Animate e.g. The child told the chicken to have a biscuit.	22 (33%)	44 (67%)	5 (24%)	16 (76%)	20 (26%)	56 (74%)	47 (29%)	116 (71%)
TOTAL/LEVEL	64 (32%)	134 (68%)	20 (32%)	43 (68%)	81 (36%)	147 (64%)	165 (34%)	324 (66%)
$\chi^2 = 4.7; DF = 6; n.s.$								

- % in brackets indicates row percentage of errors vs. correct response per sentence type

Table U

## Results of ASK Sentences: NP1 and NP2 have equal status

SENTENCE TYPE	Level1 High school		Level2 Undergrad.		Level3 Teacher Tr.		Total	
	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.
1.) Human NP x Human  e.g. The teacher asked the student to begin the lesson.	9 (14%)	57 (86%)	3 (14%)	18 (86%)	9 (12%)	67 (88%)	21 (13%)	142 (87%)
2.) Animate x Animate  e.g. The chicken asked the cat to come in.	13 (20%)	53 (80%)	2 (10%)	19 (90%)	13 (17%)	63 (83%)	28 (17%)	135 (83%)
3.) Inanimate x Inanimate  e.g. The cake asked the table to listen carefully.	19 (29%)	47 (71%)	5 (24%)	16 (76%)	31 (41%)	45 (59%)	55 (34%)	108 (66%)
TOTAL/LEVEL	41 (21%)	157 (79%)	10 (16%)	53 (84%)	53 (23%)	175 (77%)	104 (21%)	385 (79%)
$\chi^2 = 1.77; DF = 6; n.s.$								

- % in brackets indicates row percentage of errors vs. correct response per sentence type

Table V

Results of ASK Sentences: NP2 is higher on the Topicality Hierarchy [convergence]

SENTENCE TYPE	Level1 High school		Level2 Undergrad.		Level3 Teacher Tr.		Total	
	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.
1.) Inanimate x Animate e.g. The door asked the dog to come in.	24 (34%)	42 (64%)	7 (33%)	14 (67%)	16 (21%)	60 (79%)	47 (29%)	116 (71%)
2.) Inanimate x Human e.g. The ice-cream asked the girl to eat slowly.	22 (33%)	44 (67%)	7 (33%)	14 (67%)	18 (26%)	56 (74%)	49 (30%)	114 (70%)
3.) Animate x Human e.g. The pigeon asked the visitor to come in.	21 (32%)	45 (68%)	4 (19%)	17 (81%)	15 (20%)	61 (80%)	40 (25%)	123 (75%)
TOTAL/LEVEL	67 (34%)	131 (66%)	18 (29%)	45 (71%)	51 (22%)	177 (78%)	136 (28%)	353 (72%)
$\chi^2 = 7.8; DF = 6; n.s.$								

- % in brackets indicates row percentage of errors vs. correct response per sentence type



Table W

Results of ASK Sentences: NP1 is higher on the Topicality Hierarchy [conflict]

SENTENCE TYPE	Level1 High school		Level2 Undergrad.		Level3 Teacher Tr.		Total	
	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.
1.) Animate x Inanimate e.g. The dog asked the door to come in.	25 (38%)	41 (62%)	9 (43%)	12 (57%)	30 (39%)	46 (61%)	64 (39%)	99 (61%)
2.) Human x Animate e.g. The girl asked the ice-cream to eat quickly.	6 (9%)	60 (91%)	0 (0%)	21 (100%)	7 (9%)	69 (91%)	13 (8%)	150 (92%)
3.) Human x Animate e.g. The visitor asked the pigeon to come in.	23 (35%)	43 (65%)	6 (29%)	15 (71%)	20 (26%)	56 (74%)	49 (30%)	114 (70%)
TOTAL/LEVEL	54 (27%)	144 (73%)	15 (24%)	48 (76%)	57 (25%)	171 (75%)	126 (26%)	363 (74%)

$$\chi^2 = 4.03; DF = 6; n.s.$$

- % in brackets indicates row percentage of errors vs. correct response per sentence type

Table X

## Results of PROMISE Sentences: NP1 and NP2 have equal status

SENTENCE TYPE	Level1 High school		Level2 Undergrad.		Level3 Teacher Tr.		Total	
	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.
1.) Human NP x Human  e.g. John promised Bill to study hard.	39 (59%)	27 (41%)	16 (76%)	5 (24%)	40 (53%)	36 (47%)	95 (58%)	68 (42%)
2.) Animate x Animate  e.g. The dog promised the cat to leave.	25 (38%)	41 (62%)	15 (71%)	6 (29%)	38 (50%)	38 (50%)	78 (48%)	85 (52%)
3.) Inanimate x Inanimate  e.g. The flower promised the tree to stay.	36 (55%)	30 (45%)	2 (10%)	19 (90%)	45 (59%)	31 (41%)	83 (51%)	80 (49%)
TOTAL/LEVEL	100 (51%)	98 (49%)	33 (52%)	30 (48%)	123 (54%)	105 (46%)	256 (52%)	233 (48%)

$$X^2 = 11.6; DF = 6; *P < .07$$

- % in brackets indicates row percentage of errors vs. correct response per sentence type

Table Y

**Results of PROMISE Sentences: NP1 is higher on the Topicality Hierarchy [convergence]**

SENTENCE TYPE	Level1 High school		Level2 Undergrad.		Level3 Teacher Tr.		Total	
	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.
1.) Animate x Inanimate 39 - 27 (59% 41%) e.g. The lion promised the tree to leave.	17	4	43	33	99	64	(81% 19%)	(57% 43%) (61% 39%)
2.) Human x Inanimate 34 - 32 (52% 48%) e.g. The teacher promised the chair to go.	16	5	55	21	105	58	(76% 24%) (72% 28%)	(64% 36%)
3.) Human x Animate 35 - 31 (53% 47%) e.g. The man promised the cat to go.	10	11	24	52	69	94	(48% 52%) (32% 68%)	(42% 58%)
TOTAL/LEVEL	108	90	43	20	122	106	(55% 45%) (68% 32%)	(54% 46%) (273 - 216) (56% 44%)
$\chi^2 = 9.0; DF = 6; n.s.$								

- % in brackets indicates row percentage of errors vs. correct response per sentence type

Table Z

**Results of PROMISE sentences: NP2 has a higher status on the  
Topicality hierarchy**

SENTENCE TYPE	Level1 High school		Level2 Undergrad.		Level3 Teacher Tr.		Total	
	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.	NP1 Err.	NP2 Cor.
1.) Inanimate x Human  e.g. The chair promised the teacher to go.	31 (47%)	35 (53%)	13 (62%)	8 (38%)	29 (38%)	47 (62%)	73 (45%)	90 (55%)
2.) Inanimate x Animate  e.g. The tree promised the lion to leave.	33 (50%)	33 (50%)	12 (57%)	9 (43%)	29 (38%)	47 (62%)	74 (45%)	89 (55%)
3.) Animate x Human  e.g. The cat promised the man to go.	34 (52%)	32 (48%)	13 (62%)	8 (38%)	32 (42%)	44 (58%)	79 (48%)	84 (52%)
TOTAL/LEVEL	98 (49%)	100 (51%)	38 (60%)	25 (40%)	90 (39%)	138 (61%)	226 (46%)	263 (54%)
$\chi^2 = 9.55; DF = 6; n.s.$								

- % in brackets indicates row percentage of  
errors vs. correct response per sentence type

Figure A

Overall Patterns for Tell sentences

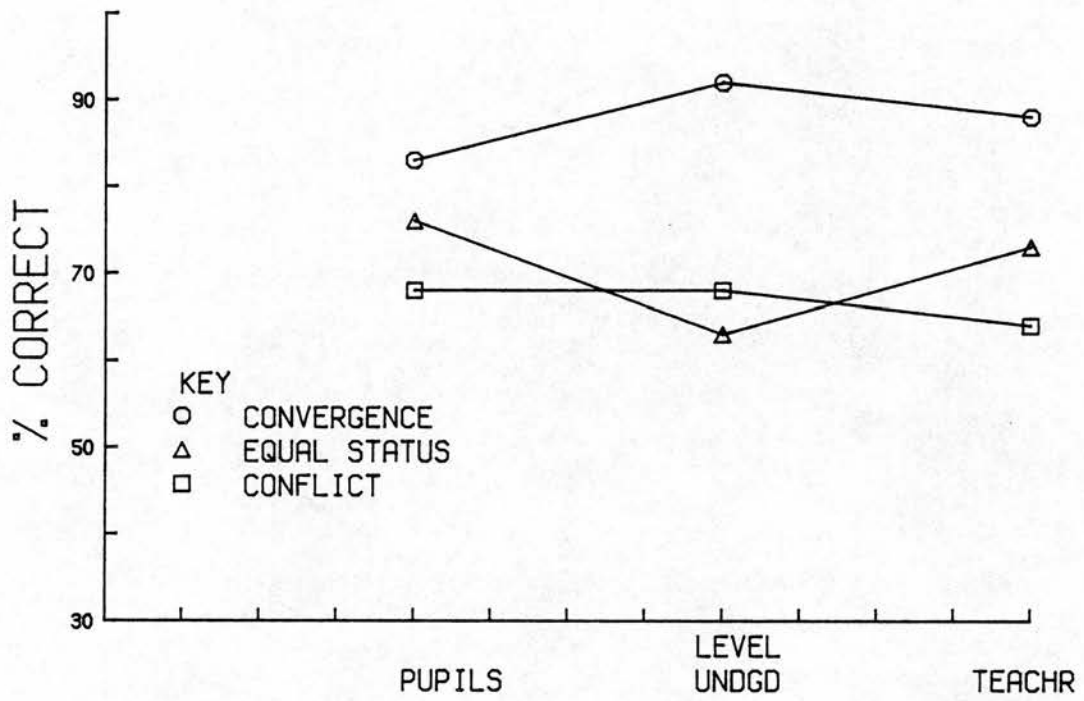


Figure B

Overall Patterns for Ask sentences

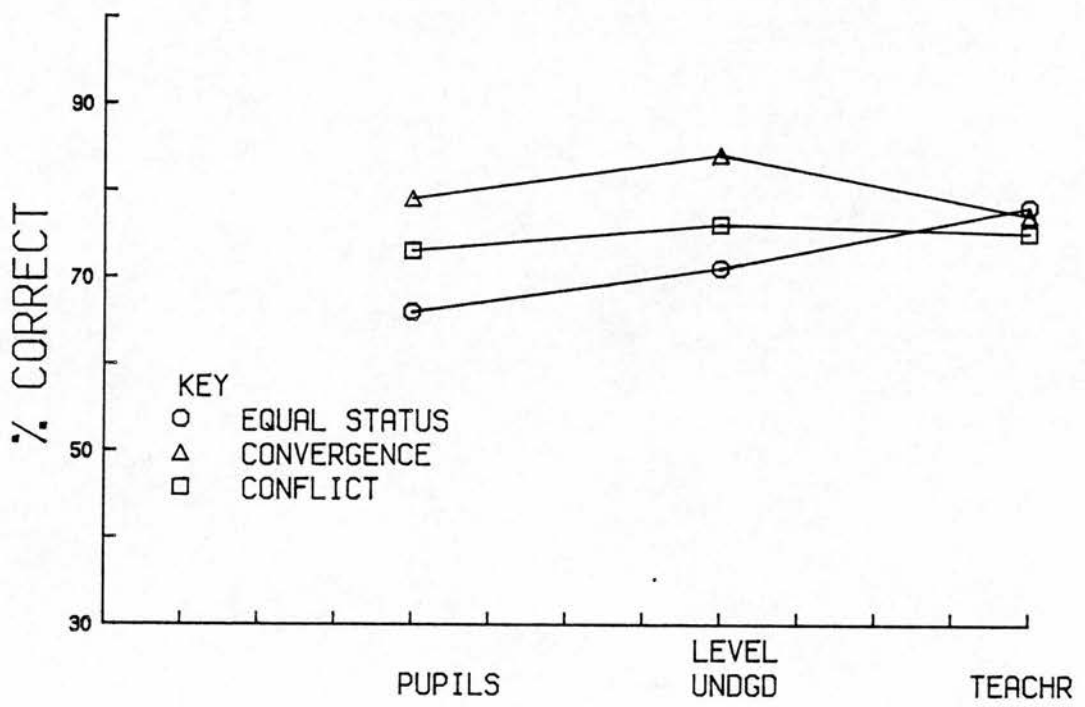
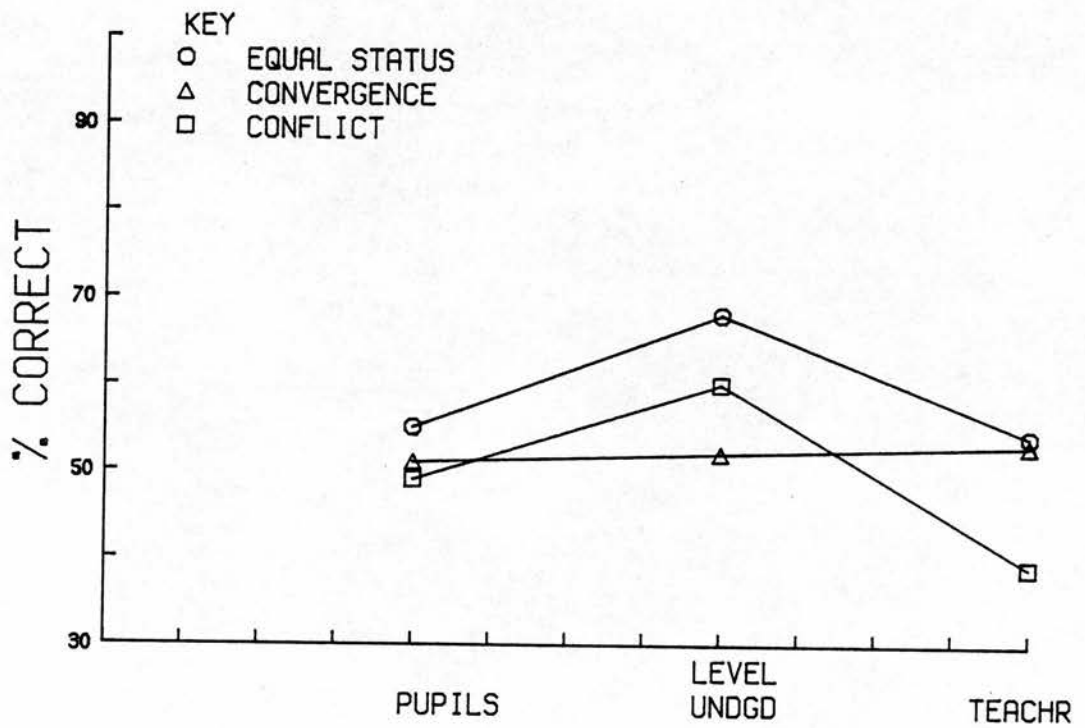


Figure C

Overall Patterns for Promise sentences



## II. Testing Instruments for the Modified Cloze Test

### Task I: A Modified Cloze Test

Full Name: \_\_\_\_\_ Date: \_\_\_\_\_

Class / Year: \_\_\_\_\_ School / College: \_\_\_\_\_

Instructions:

1. Read the following passage carefully.
2. Complete the text with the correct article: "the", "a", or "an".
3. If you think no article is needed, leave the space blank.

TIME: 20 minutes

#### Visiting YA MORO

It is 10 o'clock in 1 morning on 2 Saturday. 3 sun is shining. My sister Mbilia and I decide to go to 4 zoo, near 5 Central market in Kinshasa city. We take 7 SOTRAZ bus and get off at 8 bus stop, near 9 Mama Yemo Hospital. 10 zoo is open. I buy 11 tickets from 12 small window near 13 entrance. Mbilia and I have never visited 14 zoo before. Kinshasa zoo is 16 big one and we don't know where to begin. But 17 young boy has just noticed that we are having 18 problems. So he comes to help us.

'You want to see 19 animals? Come with me', he says. 'I am your friend and I know 20 zoo very well. I always help 21 young people because I am young too. My name is Kapwepwe. Let's begin 22 visit' says 23 the boy.

'Mbilia, I think we are very lucky today. This boy is going to serve as 24 guide for us. So we will save 25 time', I tell my sister.

'What is that?' asks Mbilia.

'Psscht' That's Ya Moro, 26 gorilla. He's 27 oldest and strongest resident in 28 zoo', says 29 boy. 'As you can see 30 gorilla is sleeping under 31 big, green mango tree. 32 gorillas like 33 mangoes and 34 bananas very much. That's why Ya Moro is sleeping under 35 tree', continues 36 boy.



'Ah! Ha!, Kapwepwe, I have 337 few bananas in my bag. Let's wake Ya Moro up and give him 38 banana', I say.

'No! You can't give him 39 banana or 40 mango. Only 41 zoo officials can do that', says Kapwepwe, 42 young guide. 'But you can sing 43 song for him. Ya Moro likes 44 songs very much, I can tell you', he continues.

'Can you sing 45 songs?' my sister Mbilia asks Kapwepwe.

'Oh! Yees! Of course!' says 46 boy. Then he begins to sing 47 exciting little song. Really, Kapwepwe has 48 very good voice. 49 song goes something like this:

Ya Moro is one of 50 biggest and 51 laziest gorillas

He eats 52 lots of bananas

And drinks 53 dirty water

From 54 small river

Ya Moro come here

If you are 55 biggest and 56 strongest come here

Stop sleeping. Come here now

Kapwepwe is here to show you how

He is 57 biggest and 58 strongest

Stop sleeping. Come here now.

Kapwepwe can beat you because he is 59 greatest

Stop sleeping Ya Moro. Come here now.

Kapwepwe sings 60 song, again and again. Then, Ya Moro hears him. Ya Moro, 61 gorilla, gets very angry. He begins to jump from one side of 62 big cage to another. He is making 63 terrible noise now. At this moment 64 zoo inspector comes to see what is happening to 65 gorilla. 66 inspector decides to give him some food, 67 water, and 68 ice-cream. Ya Moro eats 69 food. 70 gorilla likes 71 ice cream but he doesn't like 72 water. When

Ya Moro finishes eating he becomes very nice. He starts laughing, and says 'Hello!' to 73 visitors. Then my sister and I decide to move to 74 next cage.

### III. Testing Instruments for Sentence Interpretation

#### Task II: Sentence Interpretation

Full name:

Date:

Class/year of study:

School/college:

#### Instructions:

1. In this test you are asked to read every sentence carefully.
2. Then you will answer the question below ( = en dessous) that sentence. Only one answer is correct.
3. Put an (X) to show your choice.

This is an example: 'Kitoko takes his book to go to school'.

Question:- Who should go to school? 1. ( ) the book

(Qui doit aller a l'ecole?) 2. ( X ) Kitoko

The correct answer is: 'Kitoko'.

4. Now you do the same. Work as quickly as possible. Follow the example and use your own judgment.

TIME: 45 minutes

1. The boy told the girl to leave.  
Who should leave?
  1. ( ) the boy
  2. ( ) the girl
2. The man promised the cat to go.  
Who should go?
  1. ( ) the man
  2. ( ) the cat
3. The pen told the book to stop.  
Who should stop?
  1. ( ) the book
  2. ( ) the pen
4. The teacher asked the student to begin the lesson.  
Who should begin?
  1. ( ) the teacher
  2. ( ) the student
5. The door asked the dog to come in.  
Who should come in?
  1. ( ) the door
  2. ( ) the dog
6. The cat promised the man to go.

- Who should go? 1. ( ) the cat  
2. ( ) the man
7. The cake asked the table to listen carefully.  
Who should listen? 1. ( ) the table  
2. ( ) the cake
8. Mary asked Louise what to eat.  
Who should eat something? 1. ( ) Louise  
2. ( ) Mary
9. The child told the chicken to have a biscuit.  
Who has a biscuit? 1. ( ) the chicken  
2. ( ) the child
10. The soldier told the car to stop.  
Who should stop? 1. ( ) the soldier  
2. ( ) the car
11. The ice-cream asked the girl to eat slowly.  
Who should eat slowly? 1. ( ) the girl  
2. ( ) the ice-cream
12. John promised Bill to study hard.  
Who should study hard? 1. ( ) John  
2. ( ) Bill
13. The woman preferred for her friend to stay at home.  
Who should stay? 1. ( ) the woman  
2. ( ) her friend
14. The dog told the lion to wait.  
Who should wait? 1. ( ) the lion  
2. ( ) the dog
15. The dog asked the door to come in.  
Who should come in? 1. ( ) the door  
2. ( ) the dog
16. The student asked the teacher to begin the lesson.  
Who should begin? 1. ( ) the student  
2. ( ) the teacher
17. The chicken told the child to have a biscuit.  
Who has a biscuit? 1. ( ) the chicken  
2. ( ) the child
18. The chair promised the teacher to go.  
Who should go? 1. ( ) the chair  
2. ( ) the teacher
19. The chicken asked the cat to come in.  
Who should come in? 1. ( ) the cat  
2. ( ) the chicken

20. The teacher promised the chair to go.  
Who should go? 1. ( ) the chair  
2. ( ) the teacher
21. The tree promised the lion to leave.  
Who should leave? 1. ( ) the tree  
2. ( ) the lion
22. The man ordered his boss to speak slowly.  
Who should speak slowly? 1. ( ) the boss  
2. ( ) the man
23. The girl asked the ice-cream to eat quickly.  
Who should eat? 1. ( ) the ice-cream  
2. ( ) the girl
24. The car told the soldier to stop.  
Who should stop? 1. ( ) the car  
2. ( ) the soldier
26. The chicken told the house to go away.  
Who should go? 1. ( ) the house  
2. ( ) the chicken
27. The lion promised the tree to leave.  
Who should leave? 1. ( ) the lion  
2. ( ) the tree
28. The visitor asked the pigeon to come in.  
Who should come? 1. ( ) the visitor  
2. ( ) the pigeon
29. The flower promised the tree to stay.  
Who should stay? 1. ( ) the tree  
2. ( ) the flower
30. The boy hoped for his mother to receive a birthday present.  
Who should receive? 1. ( ) the boy  
2. ( ) the mother
31. Juma asked Tabu where to go.  
Who should go somewhere? 1. ( ) Juma  
2. ( ) Tabu
32. The pigeon asked the visitor to come in.  
Who should come? 1. ( ) the pigeon  
2. ( ) the visitor
33. The house told the chicken to go away.  
Who should go? 1. ( ) the house  
2. ( ) the chicken
34. The dog promised the cat to leave.  
Who should leave? 1. ( ) the cat

2. ( ) the dog

35. The barman wanted the clients to take the money.

Who should take the money?

1. ( ) the clients

2. ( ) the barman

36. The children promised the father to buy a family car.

Who should buy?

1. ( ) the children

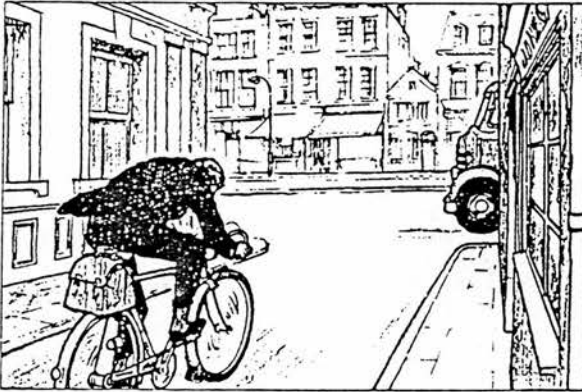
2. ( ) the father

IV. Testing Instruments for Composition (Task 3)

COMPLETE IT

Composition

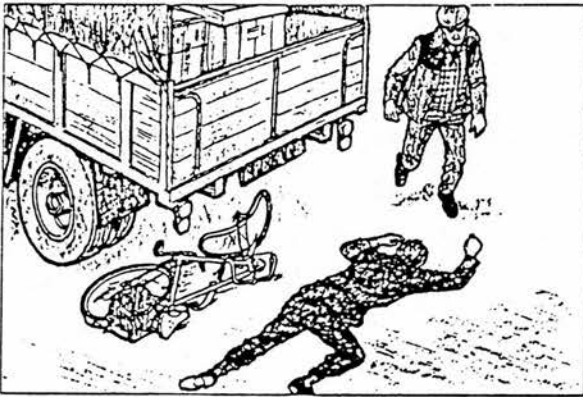
A



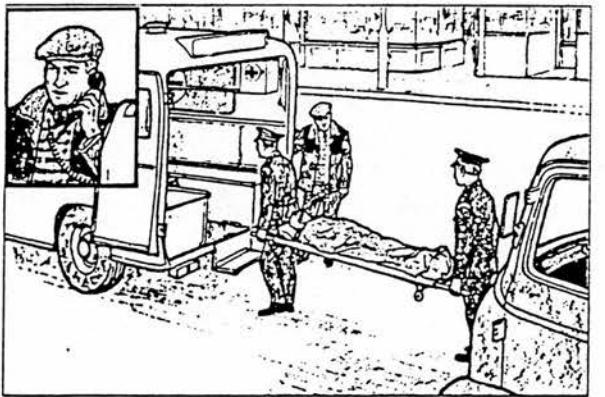
B



C



D



Question : What has just happened?

**Questions for 'Free Composition' or 'Interviews'**

1. How do you go from home to school/campus?
2. How do you get on with the girls/boys at school?
3. How many children would you like to have when you get married? Do you think it is exciting to have twins?
4. Can you describe for me your ideal wife/husband?
5. What do you think about 'permanent education', e.g. studying after graduating from college?
6. If you were a successful Zairean businessman. . . how would you use your money? Which district of Kinshasa would you choose to live in? Why?
7. Can you define for me the most interesting subject in your curriculum? (Its object, usefulness, job prospects. . .)
8. For what reasons do you sometimes go downtown?
9. Would you like to marry a girl/young man who is studying in the same class/year as yourself?
10. Do you read a lot? On what? Can you remember anything interesting that you read recently?



## V. Questionnaires for the Attitude-Motivation Survey

### I. GENERAL INFORMATION ON THE SUBJECT

Give an answer to every question.

1. Full name:
2. Year of Study/Class:
3. Name of \_\_\_ School:  
                   \_\_\_ College:  
                   \_\_\_ University:
4. Sex: ( ) M ( ) F
5. Mother tongue:
6. Which of these national languages do you speak?  
     Lingala \_\_\_  
     Swahili \_\_\_  
     Kikongo \_\_\_  
     Tshiluba \_\_\_
7. Which zone of Kinshasa do you live in?
8. How many years have you lived in Kinshasa?
9. What foreign languages do you know?
10. Which language(s) do you use:
  - 1) to write a letter to your parents:
  - 2) to write a letter to your friends:
  - 3) to write a letter to your headmaster:
  - 4) to write a letter to your English teacher:
  - 5) to apply for a scholarship from the British/U.S. Embassy:
11. Which language(s) do you speak?
  - 1) At Home
    - with your father:
    - with your mother:
    - with your brothers and sisters:
    - with your friends:
    - with your visitors:
  - 2) At school
    - in the classroom:
    - outside the classroom (playing games):
12. When (which year of study) did you start learning French? \_\_\_\_\_  
     How old were you? \_\_\_\_\_

Poorly    Fairly    Well    Very well  
Well

14. How do you speak French?  
listen to French?  
read French?  
write French?

Tick ( ) your choice

Poorly    Fairly    Well    Very well  
Well

15. How do you speak English?  
listen to English?  
read English?  
write English?

16. Do you know anyone who speaks English but who is not a student  
in your school/ college/ campus?

Yes \_\_\_\_\_ No \_\_\_\_\_

17. If yes, what is his nationality:  
occupation:

18. Do you ever have a chance to read a book, or a magazine, or  
a newspaper in English?

\_\_\_\_\_ very frequently  
\_\_\_\_\_ frequently  
\_\_\_\_\_ sometimes  
\_\_\_\_\_ rarely  
\_\_\_\_\_ never

19. Do your parents encourage you to learn

1) French:    Yes \_\_\_\_\_    No \_\_\_\_\_  
2) English:    Yes \_\_\_\_\_    No \_\_\_\_\_

13. When did you start learning English? \_\_\_\_\_  
How old were you? \_\_\_\_\_

**II. ZAIREAN LEARNERS' ATTITUDE TOWARDS ENGLISH INSTRUCTIONS:** In this questionnaire you are asked to give your opinion regarding the following statements. You may agree or disagree with them. You will show your opinion by putting the letter corresponding to your choice in the space before that sentence.

There are six alternative choices. Give only one per statement.

A = little support of the statement

B = moderate support

C = strong support

D = little disagreement

E = moderate disagreement

F = strong disagreement

1. \_\_\_\_ If the English course was removed from the curriculum it would be deeply regrettable for Zaire.
2. \_\_\_\_ English speaking countries and their people are more influential in today's world.
3. \_\_\_\_ The British people have every reason to be proud of their language and culture.
4. \_\_\_\_ Zaireans have a lot to gain from learning English.
5. \_\_\_\_ Zaireans would greatly benefit from adopting the British people's way of life.
6. \_\_\_\_ Young Zaireans should make a great effort to meet English speaking people.
7. \_\_\_\_ It would be desirable for Zairean ways to be similar to those of the Europeans.
8. \_\_\_\_ If I was given the choice, I would prefer to live in London rather than in Kinshasa.
9. \_\_\_\_ I wish there were more English speaking people than French speakers in Kinshasa.
10. \_\_\_\_ The English course would be more interesting if it was taught by an Englishman than a Zairean.
11. \_\_\_\_ All English speaking people who live in Zaire contribute to Zaire's development.

12. \_\_\_\_ The English speaking countries have produced more scientists (or inventors) than the French.
13. \_\_\_\_ Most English speaking countries (e.g. UK, USA, Nigeria) are more democratic than the French (e.g. France, Belgium, Senegal).
14. \_\_\_\_ Teaching English in Zaire is a threat to the authentic Zairean culture.
15. \_\_\_\_ In future, I would to visit the main cities where English is widely spoken (e.g. New York, London, Dallas) rather than those where French is (e.g. Paris, Brussels, Geneva).
16. \_\_\_\_ If I had a good knowledge of English, I would rather listen to the B.B.C. World Service than Radio France 'INTER'.
17. \_\_\_\_ The Voice of Zaire (the official Radio-T.V. network) should introduce programmes for learning English.
18. \_\_\_\_ I would like to marry an English-speaking person so that I could improve my knowledge of English.
19. \_\_\_\_ It is preferable to use English / French (delete one) in science and maths because it is richer than the Zairean languages (Lingala).
20. Among all the expatriates in Kinshasa, I would like to know (Put 1, 2 or 3 in the box to indicate your decreasing order of preference):

Americans  
Arabs (Lebanese)  
Belgians  
British  
French  
Japanese  
Other: Please Specify \_\_\_\_\_

**III. ZAIREAN LEARNERS' ORIENTATION INDEX** Instructions: In this questionnaire you are asked to show your opinion by putting a tick ( ) against your choice.

1. English is important for me because I need good marks in English to pass to the higher class (form).

( ) I agree  
 ( ) I slightly agree  
 ( ) I have no opinion  
 ( ) I disagree

2. English is important for me because it would enable me to have English speaking friends in future.

( ) I agree  
 ( ) I slightly agree  
 ( ) I have no opinion  
 ( ) I disagree

3. English is important for me because if you don't know English, people think you haven't got a complete education.

( ) I agree  
 ( ) I slightly agree  
 ( ) I have no opinion  
 ( ) I disagree

4. English is important for me because that would help me to understand English speaking people and their culture

( ) I agree  
 ( ) I slightly agree  
 ( ) I have no opinion  
 ( ) I disagree

5. English is important for me because it would give me greater opportunities of getting a well-paid job.

( ) I agree  
 ( ) I slightly agree  
 ( ) I have no opinion  
 ( ) I disagree

6. English is important for me because it would help me make contacts with a lot of people.

( ) I agree  
 ( ) I slightly agree  
 ( ) I have no opinion  
 ( ) I disagree

7. English is important for me because unless you have a good command of English you can't get a scholarship nowadays.

- ( ) I agree
- ( ) I slightly agree
- ( ) I have no opinion
- ( ) I disagree

8. English is important for me because it would enable me think like the British or the Americans.

- ( ) I agree
- ( ) I slightly agree
- ( ) I have no opinion
- ( ) I disagree

9. English is important because the best POP stars (like, Bob Marley, Michael Jackson, etc.) sing in English

- ( ) I agree
- ( ) I slightly agree
- ( ) I have no opinion
- ( ) I disagree

10. English is important for me because I would like to correspond with English speaking families or students for exchange of ideas /experiences

- ( ) I agree
- ( ) I slightly agree
- ( ) I have no opinion
- ( ) I disagree

**IV. ZAIREAN STUDENTS DESIRE TO LEARN ENGLISH** Instructions: In this questionnaire you are asked to put a tick ( ) against the alternative corresponding to your opinion.

1. Put a tick to indicate how much you like the English course compared to 10 other courses on your curriculum:

English is the least preferred course

English is the most preferred course

0 1 2 3 4 5 6 7 8 9 10

2. When I have homework to do in English,  
 I do it immediately before doing any other homework  
 I feel annoyed  
 I do it after I have finished with everything  
 I ask someone, e.g. a friend, to do it for me
3. During English Classes, I  
 tend to dream about other things  
 often fall asleep  
 make an effort to try and understand  
 ask a lot of questions
4. If I knew enough English, I would read short stories in English  
 very often  
 from time to time  
 perhaps, but I'm not sure  
 never
5. After I have studied English for some time  
 I feel that I want to continue studying  
 I become bored  
 I find it more and more difficult  
 I understand everything easily
6. If I had the opportunity to change the English programme and its method of teaching I would  
 keep the number of hours as it is now  
 increase the number of hours  
 decrease the number of hours  
 remove it from the curriculum
7. If I was asked to design a new curriculum, I would suggest that  
 English be taught only to those who are interested  
 English be taught to all secondary and university students  
 only spoken English be taught  
 both spoken and written English be taught
8. In my view, learning English is  
 very interesting



- of a limited interest only
- not more interesting than other subjects
- a waste of time

9. Before coming to school/English lessons,

- I always prepare my lesson
- I only prepare myself if there is homework
- I am rarely prepared
- I never prepare before

10. After I finish my studies, I will

- try to continue learning English
- not want to study English any more
- join an English club or similar organization
- see; but I have no plans now