Phonological Processes in the Acquisition of Liquid and Stop Segments in English by Anaang Speakers

Unyierie Angela Walter Idem

PhD
University of Edinburgh
1994

DECLARATION

I hereby declare that this thesis has been composed by me and is entirely my own work.

Unyierie Angela Walter Idem

ABSTRACT

Research has shown that the second language (L2) learner's phonological system is shaped by a number of processes which interact in various ways to produce a variable interlanguage system. This study examines the extent to which transfer, developmental and universal processes operate in the acquisition of liquid (/r,l/) and stop (p, b, t, d, k, g/) segments in English by native speakers of Anaang.

A number of different elicitation tasks are used: word lists, sentence lists, texts and interview conversation. Liquid segments are tested in four phonetic environments across tasks - initial, cluster, medial and final - and stop segments in the final position across the same range of tasks. The data are collected from Anaang speakers of different proficiency levels in L2 English: low, lower intermediate, upper intermediate and advanced. Data are also collected from native speakers of English for purposes of comparison.

The results indicate that while transfer remains the predominant process, it interacts constantly with developmental and universal processes. Furthermore, the manifestation or non manifestation of a particular process and the degree of its influence on the phenomena tested are determined by the linguistic context, the speech situation (casual or careful) and the learners'

level of proficiency in English. These variables interact in complex ways in that the differences observed among proficiency groups are mediated by the phonetic environments of the segments and the formality of the tasks administered. There is need therefore for an adequate theory of second language phonology which will account for such variable phenomena resulting from the interaction of different processes.

ACKNOWLEDGEMENTS

This work could not have been completed without the co-operation and support of a great many people. My gratitude goes, first and foremost, to my supervisor Dr. Antonella Sorace for her unfailing guidance, encouragement and patience throughout the whole study. I am also grateful to Mr. Anthony Howatt who offered me invaluable advice at the different stages of this work. Dr. Dan Robertson was most helpful with statistical analyses, and to him I am deeply indebted.

My sincere thanks go to the following colleagues and friends whose assistance and co-operation made the data-gathering process less tedious. In Nigeria they include Imo Eshiett, Isidor Ekpe, Anietie Idok, Joseph Umo, Rosemary Edet, Nyeme Chuku and Okon Sam. I am particularly grateful to them for assisting in recruiting participants for the study, and for the much useful linguistic information they provided on the Anaang language. I have not forgotten my cousin, Anietie Marcus Akpan, who assisted in the same capacity. In Edinburgh Bruce Connell, David Hill, Kayoko Enomoto, Richard Warner and Yuko Kondo offered invaluable advice, suggestions and practical help. I am also grateful to Cedric MacMartin and Irene Macleod at the Phonetics Laboratory Edinburgh University for technical assistance.

To the students who participated in the study and made the completion of this work a reality, I owe my profound gratitude.

I would also like to express my gratitude to the following: Charles Mann and Barineme Fakae for their friendship, moral support and great sense of humour: Eno Urua, Josephine Munthali, Matthew Olaniyan, Brigid Asuquo, Uwem and Ernest Eteng, Ify and Hubert Ezirim, Ese and Osato Osaghae, Florence and Kingsley Larbi for their kindness and encouragement. Mr. Ekanem Udok will be remembered for being fatherly and protective.

I am greatly indebted to the Association of Commonwealth Universities/British Council for sponsoring my studies in the United Kingdom, and to the Leche Trust for financial support towards part of the research cost. My thanks go to the Ministry of Education, Uyo for granting me study leave to further my studies.

Finally, I would like to thank my parents, Mr. and Mrs. Idem, my brothers, Uko, Kufre and Saade, and my sisters, Anietie and Irene for their support, morally and financially, and their confidence in me.

This thesis is dedicated to my mother

for standing by me all the way.

NOTES ON SYMBOLS AND TRANSCRIPTIONS

The phonetic symbols used in this work conform to the conventions of the International Phonetic Association (IPA). Nevertheless, a number of authors quoted in the study have used symbols of other transcription systems (e.g. the Bloch and Trager system, and the American system). Such symbols have been retained for reasons of authenticity.

A common difference noted between American and IPA systems, for example, is the use in the American transcription, of the wedge diacritic (ˇ) on symbols for palato-alveolar fricatives ([š], [ž]), palato-alveolar affricates ([č], [ǰ]), and liquids with a flaplike or taplike quality ([ř], [ǐ]). Compare the IPA symbols [ʃ], [ʒ], [tʃ], [dʒ], [ɾ] and [ɻ] respectively.

The wedge diacritic has its use in the IPA system as a transcription for a rising tone. We follow this convention in the present study. Other tone markings are:

(ˊ) high tone

(ˋ) low tone

(ˆ) falling tone

(ꜜ) downstep tone

The tones are used in the transcription of examples from Anaang, the orthography of which is given below. The

phonemic and phonetic transcriptions represented by the symbols / / and [ ] respectively are also indicated.

The Orthography of Anaang

Letters	Gloss	Phonemic transcription	Phonetic transcription
a	álàn (oil)	/a/	[a]
b	bén (take)	/b/	[b]
ch	chóp (be lost)	/t ʃ/	[tʃ]
d	dép (buy)	/d/	[d]
e	bét (wait)	/e/	[e]
f	fìré (forget)	/f/	[f]
gh	fèghé (run)	/ɣ/	[ɣ]/[R]
gw	gwèt (write)	/gʷ/	[gʷ]
i	sín (put)	/i/	[i]
j	jém (look for)	/dʒ/	[dʒ]
k	kǎ (go)	/k/	[k]
kp	ékpè (lion)	/k͡p/	[k͡p]
l	ílìm (stream)	/l/	[l]

m	ímá (love)	/m/	[m]
n	nék (dance)	/n/	[n]
ñ	saña (walk)	/ɲ/	[ɲ]
ñw	ñwéék (breathe)	/ɲ/	[ɲʷ]
ny	nyàm (sell)	/ɲ/	[ɲ]
o	bòm (break)	/o/	[o]
ọ	bọ́k (make soup)	/ɔ/	[ɔ]
p*	fọ́p (roast)	/p/	[p-]
r	íbòró (answer)	/r/**	[ɾ]
s	sìp (be small)	/s/	[s]
t	tèm (cook)	/t/	[t]
u	úfòk (house)	/u/	[u]
u̱	dú̱k (enter)	/ʉ/	[ʉ]
w	wìré (play)	/w/	[w]
y	kpéyé (plead)	/j/	[j]

* Occurs mainly in word final position. Note the use of the symbol (-) to indicate the non release of a stop.

** /r/ is phonemic only in some dialects of Anaang (e.g. Abak), and is often represented by /l/ in others (e.g. Ukanafun)

CONTENTS

CHAPTER NINE: ANALYZING THE PLOSIVE DATA

FIGURES

TABLES

APPENDICES

CHAPTER ONE: INTRODUCTION

1.1 Background to the study

Until recently the study of interlanguage (IL) phonology remained a largely neglected area within second language acquisition (SLA) research. One reason for the dearth of studies in this domain was the assumption that the investigation of the second language (L2) learner's phonological system would offer no useful or interesting insights into the nature of the SLA process, since all phonological errors were thought to derive directly from first language (L1) interference (Ioup and Weinberger, 1987; Tarone, 1978). There was also, as Tarone suggests, the general conviction on the part of SLA researchers, teachers and students that the pronunciation of a second language was less important than its grammar and vocabulary systems.

However, the past decade and a half have witnessed a resurgence of interest in the phonological aspects of SLA research. This can be seen in the appearance of anthologies of articles devoted to this domain (Ioup and Weinberger, 1987: James and Leather, 1987; Leather and James, 1990). Some phonological papers have also appeared in Gass and Schachter (1983), Gass and Madden (1985), Kellerman and Sharwood Smith (1986). These and other studies have shown that transfer is not the simplistic phenomenon it was once thought to be. More importantly,

that the L2 learner's phonological system is a complex system resulting from the interplay of several factors of which transfer is only one. Others are developmental factors (Flege, 1980; Flege and Davidian, 1984; Major, 1987a, 1987b; Hecht and Mulford, 1987; Nemser, 1971; Piper, 1984), language universal constraints (Anderson, 1987; Broselow, 1987: Broselow and Finer, 1991: Dreasher and Anderson-Hsieh, 1990; Eckman, 1977, 1987; Tarone, 1980), as well as sociolinguistic and stylistic determinants (Beebe, 1987a; Dickerson and Dickerson, 1977; Schmidt, 1987; Tarone, 1983). There are also non structural factors which include age (Neufeld, 1980; Scovel, 1969; Seliger et al., 1975), individual differences (Guiora et al., 1975) and level of L2 proficiency (Major, 1986a, 1987a; Taylor, 1975).

These factors, together, account for the "foreignness" which often marks second language speech. Major (1986a:53) has defined foreign accent as "a deviation in pronunciation from the norms of native speakers of the language." SLA researchers have been concerned with "how", "when" and "why" these factors cause such deviations, as well as the nature of interactions between factors. In other words, they have been concerned with the description as well as explanation of the psycholinguistic processes of L2 phonological acquisition. Hence the plethora of SLA theories, some of which imbibe the principles of

linguistic, and in this case phonological theories (see Chapter Three).

However, this does not imply a purely theoretical approach to the study of L2 phonology, as a substantial part of the research has been motivated primarily by pedagogical concerns. The era of classical Contrastive Analysis is noted for pedagogically-oriented studies which aimed at making language teaching more effective. Yet, while contemporary studies of SLA do have implications for language teaching, there is growing caution about the conclusions to be drawn from such findings. As Odlin (1989) argues, "this caution reflects a growing awareness of the complexity of transfer and other topics in second language research" (p. 157).

1.2 Research aims and questions

Studies of interlanguage phonology and the processes which operate therein have been carried out across many language contact situations. To our knowledge, there are no such studies on a language contact situation involving English and Anaang (a Benue-Congo language spoken in parts of south eastern Nigeria). The main concern of this study, therefore, is the investigation of the processes that shape the IL English phonology of Anaang speakers. We are particularly interested in the roles of transfer, developmental and universal processes in the acquisition of liquid and stop speech sounds in L2 English. Acquisition is examined here

from two perspectives: production and perception. Liquid segments are considered from the perspective of production and perception, and stops from the point of view of production alone. In the case of the former the intention is to investigate the nature of the relationship between the productive and perceptual mastery of L2 sounds.

Given the above considerations the following main research questions arise:

(a) To what extent does native language transfer affect the acquisition of liquid and stop speech sounds in L2 English?

(b) To what extent do developmental factors account for the acquisition of liquid and stop speech sounds in L2 English?

(c) To what extent do language universal factors influence the acquisition of liquid and stop speech sounds in L2 English?

(d) What is the relationship between the production and perception of liquid contrasts in L2 English?

In order to answer these questions the following variables are taken into consideration: level of L2 proficiency, linguistic contexts of the segments and communication tasks. These factors serve as constraints on phonological processes in L2 speech acquisition.

1.3 Structure of the Thesis

The thesis consists of eleven chapters. This introductory part (**Chapter One**) has provided a brief background to the issues which the present study intends to address. The research aims and questions which arise therefrom have also been presented. The rest of the thesis will present the literature review, design and methodology, data analysis and discussion of the results obtained from the investigation.

The review of related literature will be undertaken in Chapters Two and Three. **Chapter Two** will examine research trends in the area of interlanguage phonology. This will involve a discussion of the major findings of research studies (in the light of the factors identified above), their methodological approaches and analyses. **Chapter Three** will be concerned with the theoretical dimension of phonological research. We shall begin by examining general theories of phonology, and in particular their explanatory potential for phonological acquisition. While the theories focus on L1 acquisition, their applicability to the L2 situation will be discussed. It will be argued that the theoretical proposals account for L2 phonological acquisition, but with some limitations.

The discussions presented in Chapters Two and Three will serve to highlight the questions raised above, and consequently lead to the formulation of hypotheses

for the study. This will form the main focus of **Chapter Four.**

In **Chapter Five** the language situation in Nigeria will be examined with a view to describing the setting in which the participants in this study have acquired their L2. The discussion reveals a highly multilingual society, with English playing the leading role against a background of Nigerian languages. The impact of these languages on the L2 acquisition of English is evident in the emerging though "stigmatised" model of Nigerian English.

Chapter Six will present the description of the specific linguistic structures to be investigated. The rationale for the selection of liquids and stop speech sounds will be given. A basic contrastive analysis approach will be adopted which will entail a comparison of the said structures in English and Anaang.

Chapter Seven will focus on the experimental design of the study as well as the procedures followed in data collection. This will include a description of the data collection instruments, the manner of administration, and the subjects to whom they were administered during the fieldstudy.

The analysis of the data obtained from administering the said instruments will be undertaken in **Chapters Eight and Nine.** The former will deal with the liquid data, and the latter with the plosive (or stop) data. The analysis will consist of descriptive as well as

inferential statistics (where the hypotheses formulated in **Chapter Four** will be tested).

Chapter Ten will discuss the results of the study. It will be shown that transfer, developmental and universal factors operate in different dimensions in L2 acquisition, and that they interrelate in different ways both synchronically and diachronically. The manifestation or non manifestation of a particular process and the degree of its influence on the phenomena tested will be shown to vary with variables such as linguistic context, the speech situation (casual or formal) and the level of L2 proficiency. On the whole, transfer will be seen to predominate over other factors.

Chapter Eleven will conclude the study, and recommendations for further research will also be made.

CHAPTER TWO: RESEARCH IN INTERLANGUAGE PHONOLOGY

2.1 Introduction

This chapter examines research trends in the area of interlanguage (IL) phonology, focusing on major findings of research studies. The chapter is divided into six sections. The first section deals with the influence of L1 structure on L2 phonological acquisition. The evolution of the concept of language transfer from the classical contrastivist era is considered, as well as the conditions which trigger transfer. In the second section, we examine the effects of developmental factors on the acquisition process. The third section deals with language universal constraints which operate to shape interlanguage phonology. While it is argued that developmental and universal factors may affect L2 acquisition independently of language transfer, we are more interested in how these factors interact with language transfer constraints. The fourth section focuses on the interrelationship between the production and perception of L2 speech sounds. The notion of interlanguage variability which takes into account structural as well as sociolinguistic factors is dealt with in section five. Finally, in section six we consider the phenomenon of L2 fossilization.

2.2 The influence of L1 structure

There is a general consensus among many researchers that L2 acquisition is in one way or another influenced by the structure of the L1 (Altenberg and Vago, 1987; Gass and Selinker, 1983; Koo, 1972: Tarone, 1980). In other words, the learner in attempting to master or communicate in the L2, makes use of the L1 knowledge. This phenomenon has been generally referred to as "language transfer" or "interference".

However, the use of the term "transfer" to describe the interrelationship between the L1 and L2 has been a source of controversy. Some scholars have drawn attention to the theory-specific nature of the term, claiming its applicability within the behaviourist framework only (Corder, 1983; Gundel and Tarone, 1983: Kellerman and Sharwood Smith, 1986). The behaviourists' notion of the phenomenon as involving the simplistic transfer of surface patterns from the L1 to L2 is questioned (see section 2.2.1 of this study). The danger of using such a technical term, according to Corder (1983), is that it constrains one's freedom of thinking about the role of L1 in L2 acquisition, which as recent research has shown, includes non-obvious effects such as avoidance, hypercorrection, borrowing and misinterpretation. In place of "transfer" therefore, the following terms, among others, have been suggested: "cross-linguistic influence" (Kellerman and Sharwood Smith, 1986), "mother tongue influence" (Corder,1983) and

"L1 influence" (Gundel and Tarone, 1983). These terms cover obvious as well as non obvious effects of L1 structure on L2.

While we do not dispute the logic behind such arguments, we however consider the term "transfer" adequate enough to describe transfer *per se* as well as other L1-based phenomena. For the purpose of this study therefore, "transfer" will refer to the effect (obvious or non obvious) of prior L1 knowledge on L2 acquisition. At the same time, we shall continue to use "transfer" interchangeably with the other terms that have been suggested.

Interestingly, the general belief is that language transfer predominates more at the phonological level than at the syntactic or semantic levels of language (Beebe, 1987b; Broselow, 1988; Karpf et al., 1980; Ioup, 1984; Scovel, 1976). Karpf et al. (1980:195) attribute this to the fact that morphological-syntactic strings generated in the L2 are filtered through the phonological redundancy rules of L1. This happens as a result of the learners' high level of concentration on other components of speech during communication, i.e. on syntactic organization, choice of lexical items and inflectional morphemes.

A physiological explanation along the lines of Lenneberg's Critical Period Hypothesis (CPH) suggests that cerebral lateralization for language functions at the onset of puberty, affects the acquisition of L2

phonology more than syntax and semantics (Lenneberg, 1967: Scovel, 1969). As a result, post-pubescent learners "never seem capable of ridding themselves of foreign accent" (Scovel, 1969:245), whereas child learners successfully overcome any traces of such accent. Those who disagree with this view argue that foreign accents have been detected in children's speech (Valette, 1964), while some adults have been shown to acquire L2 phonology without accent (Neufeld, 1980).

Perhaps, transfer only appears to predominate in phonology because of easier detectability at this level of language than at others. Thompson (1991:178) has pointed out that "of all aspects of human language, pronunciation is, perhaps, the most immediately observable". Ioup (1984) showed that native speakers of English were able to discern foreign accents in non-native speakers, based on phonological rather than syntactic cues. A simple explanation for this could be that at the syntactic level, developmental errors were mixed with any transfer errors that might have occurred. Therefore, many syntactic errors were categorized as developmental, while most phonological errors were attributed to L1 transfer.

While the role of L1 in L2 acquisition at different language levels remains debatable (Broselow and Finer, 1991; Tarone, 1980), there is no denying that language transfer has been the major pre-occupation of linguists and SLA researchers for a considerable period

of time. As early as 1899, Sweet (cited in Ringbom, 1987) noted that L1 influence on L2 was likely to occur when the two languages were similar rather than when they were different. The same view was expressed by Palmer (also cited in Ringbom, 1987). Trubetzkoy (1939), writing on phonology said:

> The sounds of the foreign language receive an incorrect phonological interpretation since they are strained through the "phonological sieve" of one's mother tongue (pp. 51-52).

These early intuitions about L1-L2 relationship, coupled with major trends in the fields of linguistics and psychology in the early part of the 20th century were to form the basis for the Contrastive Analysis Hypothesis (CAH) of which Fries (1945) and Lado (1957) have been the main advocates.

2.2.1 Contrastive Analysis

Underlying the contrastive analysis approach was the assumption that L2 acquisition took place through the simple transfer of L1 surface features into L2. Thus,

> individuals tend to transfer forms and meanings, and the distribution of forms and meanings of their native language and culture to the foreign language and culture - both productively when attempting to speak the language and to act in the culture, and receptively when attempting to grasp and understand the language and the culture as practised by natives (Lado, 1957:2).

Transfer was either positive or negative, determined by inherent similarities and differences between the two languages. These similarities and differences were identified on the basis of a contrastive comparison of the two languages.

Such a comparison provided the basis for predicting areas of ease and difficulty in L2 learning. Where there were similarities between the L1 and L2, learning was facilitated through positive transfer, while differences between the two languages resulted in learning difficulties and subsequently in negative transfer (or proactive inhibition). Weinreich (1953:1) asserted that

> the greater the difference between the systems, i.e. the more numerous the mutually exclusive forms and patterns in each, the greater is the learning problem and potential areas of interference.

Different learning problems, and therefore different interference effects were predicted for L2 learners of different native language backgrounds. Thus, Japanese, Spanish, Thai and Tagalog native speakers learning English would have different learning and transfer problems arising from differences between their native languages and English.

The contrastive analysis approach to transfer was based on the structuralist-cum-behaviourist view of

language as the sum of its smallest parts and components (sounds, words, sentences), and of language learning as the acquisition of these discrete units (Brown, 1980). A basic unit of behaviour was the habit, which was the result of externally determined conditions. Language acquisition was, thus, a product of habit-formation (Fries, 1945; Seliger, 1988; Skinner, 1957). Habits were constructed through the repeated association between some stimulus (S) and some response (R), which would become bonded and automatized through reinforcement (Horman, 1971; Paivio and Begg, 1981).

L2 acquisition was thought to operate according to the same principles of habit-formation. Learning an L2 involved acquiring a new set of habits wherever the S-R links of the L2 differed from those of the L1, i.e. by overcoming and replacing L1 habits which were seen to be interfering with acquisition of the new set (Carroll, 1968). Where there were similarities in S-R links between the two languages there was no need for the learner to overcome proactive inhibition.

In the domain of phonology/phonetics, transfer was said to affect L1 phonemes, their phonetic realizations (variants), stress, rhythm and intonation patterns as well as the distribution of these units. Difficulty would arise in the absence of any overlapping in the realization, structuring and distribution of these units between the two languages (Lado 1957). However, the most difficulty would occur where there is partial overlap,

that is where

> part of a phoneme in the native language can pass as a separate phoneme in the foreign language, and parts of the same native language phoneme pass as a different phoneme in the foreign language (1957:15).

Weinreich (1953) considers phone substitution to be the most difficult type of interference i.e. the substitution of phonemes that are identically defined in the two languages (e.g. Romansch /b/ and Schwyzertutsh /B/), but whose normal pronunciation differs. /b/ is usually voiced, while /B/ is voiced only occasionally (p. 19). He identifies other types of interference which include under-differentiation of phonemes (i.e. the confusion between a pair of sounds of the L2 whose counterparts are not distinguished in the L1); over-differentiation of phonemes (i.e. the imposition of phonemic distinctions from L1 on L2 sounds where they are not required); and reinterpretation of distinctions (i.e. the distinguishing of phonemes of the L2 system by features which in that system are redundant, but which are relevant in the L1) (1953:18).

Lado is responsible for what has been termed the "strong version" of the Contrastive Analysis Hypothesis (CAH), given its highly predictive nature (Wardhaugh, 1970). The popularity of this hypothesis produced a series of contrastive phonological studies which were

largely pedagogically oriented. The aim was to discover and predict learning difficulties for the purpose of making L2 teaching more effective. They include Sumukti (1958) who compared Javanese and Sundanese with English; Dalbor (1959/60) who compared Spanish with English; Malick (1957), Lehn and Sleger (1960), Yarmohammadi (1969) who compared Arabic with English; Parish (1963) who compared Burmese with English; Samarajiwa and Abeyskera (1964) who compared Sinhalese with English; and Dunstan (1966) who compared Igbo, Hausa and Yoruba with English.

Perhaps the most comprehensive attempt at contrastive analysis was made by Stockwell and Bowen (1965) who carried out a detailed comparison of Spanish and English phonology. In addition to comparing the surface structures of the two languages, they established a hierarchy of difficulties based on structural as well as semantic and functional similarities/differences between Spanish and English. Accordingly, the easiest forms are those which correspond structurally, semantically and functionally in the two languages, while the most difficult are those which lack such correspondences.

Another comprehensive attempt at contrastive analysis is Moulton's (1962) classification of errors made by American English learners of German. Interestingly, his study differs from Stockwell and Bowen's in that in addition to phonemics he also

considers the phonetic aspects of pronunciation, for "many of the difficulties which our students have with German pronunciation are of a phonetic rather than of a phonemic nature" (cited in Brière, 1968:19; see also Beebe, 1987b). In other words, many pronunciation errors relate to the physical properties of speech sounds, which may be articulatory, acoustic or auditory, rather than to the patterning and functional behaviour of such speech sounds in language as a communicative system. Phonetics is therefore concerned with speech sounds and their physical features, while phonology is concerned with the organization of speech sounds in a language and the part they play in manifesting the meaningful distinctions of the language (O'Connor, 1973:17).

When the predictions of the strong version of the CAH were subjected to empirical tests (Brière, 1968; Nemser, 1971) serious flaws were discovered. Its validity was questioned both on the grounds of under - and over prediction (Odlin, 1989; Kellerman, 1984; Scarione, 1970; Wardhaugh, 1970; Whitman and Jackson, 1972). Not only did it predict errors that did not materialize, it failed to predict errors that actually occurred in learners' performance data. This does not, however, imply that some predictions were not borne out, as many studies have shown that they were.

Nemser's (1971) study is a typical example of how classical phonological CA predictions can sometimes lead to correct and sometimes to incorrect results. He

investigated the perception and production of English stops and interdentals by native speakers of Hungarian who had a limited knowledge of English. Based on a CA of the two languages, he predicted that Hungarian speakers would (a) identify English stops with corresponding Hungarian stops, and (b) identify English interdentals with Hungarian labial fricatives, apical stops or grove sibilants.

These predictions were borne out for stops, but the results were mixed for interdentals. There was asymmetry in the perception and production of interdentals. While the voiced /ð/ and voiceless /θ/ interdental fricatives were perceived as labial fricatives /v/ and /f/ respectively, they were produced as apical stops /D/ and /T/ respectively. What Nemser's CA also failed to predict was that the Hungarian speakers would produce a range of approximative sounds or "phoneme blends" which were assignable neither to English nor to Hungarian. These results led to the conclusions that CA was incapable of accounting for the complex interference patterns observed, the occurrence of approximative sounds and sequences, as well as the asymmetry between perception and production. The study showed that perceptual and productive mechanisms are independent and not isomorphic as claimed by Lado and others (see section 2.6 of this study). James (1971) in the attempt to exculpate CA from such criticisms stressed that CA did not claim to predict all errors, neither did it claim

linguistic omniscience about which "choices" speakers would make.

Other arguments against CA focused on the linguistic-psychology overlap, whereby difficulty was equated with linguistic divergence. It is argued that linguistic and psychological parameters do not overlap (Brière, 1968), given that differences can only be identified by contrasting formal features, while difficulty demands psychological considerations (Ellis, 1985a). Thus, one cannot rely solely upon the analysis of a linguistic product to yield meaningful insight into a psycholinguistic process, i.e. L2 learning. This argument directly follows from evidence that differences between languages do not necessarily lead to difficulty. On the contrary, difficulty, and therefore interference, is more likely to occur where there is some similarity between the L1 and L2 (Lee, 1968; Koutsoudas and Koutsoudas, 1962; Littlewood, 1984; Oller and Ziahosseiny, 1970; cf. Sweet, 1899 cited on page 12). Where marked differences exist between the two languages, there is no basis for comparison, and therefore no basis for transfer. Partial similarities, on the contrary may, bring about confusion and interference.

It can be argued that there is little difference between this argument and that put forward by Lado. What is termed "partial similarities" here constitutes "partial differences" from Lado's point of view. We assume that partial similarity implies partial difference

and vice versa. Lado recognizes that certain phonemes which are tentatively accepted as "similar" in the two languages may have phonetically different and context-sensitive variants, and therefore cause difficulty. He stresses that this is by far the most difficult kind of interference problem to overcome.

English and Spanish, for example, share the phoneme /d/, but Spanish has two realizations of the phoneme: a stop variant [d] and a fricative variant [ð] which occurs intervocalically and after /r/. Thus, while the English phoneme /d/ does not constitute a pronunciation problem as a total unit, it is a phonemic problem in the said contexts. Whether these predictions are borne out in experimental data or not is another matter. What is worth noting is that Lado, despite his critics claim, is aware of the role of partial similarity (in other words partial difference) in interference.

The critics have also argued that language transfer is not the only cause of errors. Other causes include transfer of training (i.e. teaching methods), learners' knowledge of L2, age, aptitude, motivation, and overgeneralization (Baird, 1967; Carroll, 1968; Jenner, 1976; Lee, 1968: Odlin, 1989). James (1971) has again stressed that CA never conceived of interference as the only source, but as the chief source of errors. Moreso, CA was concerned with purely linguistic errors and not errors due to extra-linguistic factors. According to

Fisiak (1980), such extra-linguistic factors can only be accounted for by the theory of language errors.

Perhaps, the major criticisms against CA centred on its psychological as well as its linguistic basis. The principles of taxonomic phonology were considered to be inadequate for purposes of CA (Kohler, 1971). Structural linguistic theory based on surface characteristics of language (phonemes, allophonic variations and distributions of phonemes and allophones) did not provide adequate and complete descriptions of language on which CA's could be based. Moreover, there were no explicit criteria for establishing equivalences of structures of different languages. Even the generative phonology theory which was later incorporated into CA (Di Pietro, 1968) equally proved inadequate in that it did not meet the requirements of descriptive as well as explanatory adequacy (Scarione, 1970; Wardhaugh, 1970).

Zobl (1980) points out that the problem was not with the structuralist theory, but with its interpretation. He argued that CA never exploited the structuralist principle that "the receiving language (i.e. the L2) must contain certain biases or...innovation possibilities in order for transfer to make structural inroads" (p. 46). According to the structuralists, areas of the L2 susceptible to L1 influence are those exhibiting ambiguity, instability or irregularity. Whitman (1970) posits that CA failed because of the lack

of procedural objectivity on the part of the researchers. As a result, findings differed from study to study.

Chomsky (1959) condemned the behaviourist approach to language learning, arguing that language learning does not consist of a set of habits formed through S-R associations, imitation and reinforcement, but involves a process by which the learner creatively builds up knowledge of a particular language through exposure, using his/her genetically endowed language capacity. This the learner does through hypothesis formation and testing against input received.1 The main emphasis in this mentalist approach to language learning is the active contribution of the learner to the learning process.

These criticisms paved way for a weak version of the CAH which seeks not to predict, but to account for difficulties in L2 acquisition on the basis of interference from the native language. Catford (1968) considers the most important role of CA to be explanatory rather than predictive. Wardhaugh (1970) and Lee (1968) have pointed out that CA cannot be used in its predictive, but explanatory form. Gass (1980) has however argued that the weak version is of no scientific value, since it does not explain the origin of non-L1 errors.

Chomsky's mentalist approach to language acquisition was exploited by proponents of the Creative Construction Hypothesis (CCH), who argued that L2

---

1 A revised version of this proposal in the form of Universal Grammar (UG) is discussed in Chapter Three.

learners do not rely on L1 knowledge, but on universal cognitive mechanisms which guide their discovery of the rules governing the language they are dealing with (Dulay and Burt, 1974). In other words, these cognitive mechanisms operate directly on the L2 as an independent system, in much the same way as in L1 acquisition. L1 transfer therefore plays no real part in language learning. Only in the domain of phonological acquisition is the role of transfer recognized. Even then, phonological transfer is still treated as a behaviourist notion, involving the simple "transfer of motor skills" and therefore lacking creativity. However, research in SLA within the last decade and a half has emphasized the creative nature of language transfer, thereby providing a broader and more sophisticated view of the phenomenon.

2.2.2 Language transfer revisited

Underlying new approaches to transfer, therefore, is the view of the phenomenon as a cognitive strategy, which the learner adopts, consciously or unconsciously, in processing L2 data in an attempt to master, comprehend and express meaning in it (Brown, 1980; Ellis, 1985a; Littlewood, 1984; Kellerman, 1977; Mclaughlin, 1987; Selinker, 1972; Sharwood Smith, 1979). Selinker identifies transfer as one of the five central processes which underlie SLA.

Brown (1980) identifies two types of strategies: (a) learning strategy "a method of perceiving and storing particular items for recall", and (b) a communication strategy, "a method of encoding or expressing meaning in a language" (p.83). In the same vein, Faerch and Kasper (1987) distinguish between transfer as a learning strategy and as a communication strategy. The former is diachronic, "relative to a particular instance of speech production/reception, exhibiting (positive or negative) L1 influence at the product level" (p.112). The latter is synchronic in that it takes place at the moment of speech production or reception.

As a learning procedure, transfer is used in the learner's attempt to establish hypotheses about L2 rules and items. Corder (1971) points out that the L1 provides a rather rich and specific set of hypotheses which learners use. At the same time, it constrains the hypotheses that are possible about the L2 (Schachter, 1983). The learner therefore restricts his/her hypotheses to those he/her finds easier to construct on the basis of L1 knowledge.

As a communication procedure, transfer involves either the "activation of L1 knowledge in the establishment of an IL speech plan by means of which the learner seeks to realize a communication goal" or the interpretation of incoming L2 utterances on the basis of L1 knowledge (Faerch and Kasper, 1987b:113). Two kinds of knowledge are reactivated: (a) declarative knowledge - a

language user's underlying knowledge about linguistic structure, and (b) procedural knowledge - knowledge of retrieval procedures for declarative knowledge. Transfer can occur at either or both levels of knowledge, and is particularly sensitive to factors of attention and automatization (Kohn, 1986).

Transfer may therefore act as a facilitator or inhibitor of the rate with which the L2 is acquired (Kellerman, 1984). Corder (1978) observes that where the L1 and L2 are similar, transfer will cause the learner to proceed more rapidly along the "universal route".2 By implication, where there are differences, the L1 will cause the learner to proceed less rapidly along the same route, or not at all, in which case there will be fossilization. Major (1987a, 1987c) in his studies of Brazilian Portuguese (BP) learners of English has shown that similar phenomena between the L1 and L2 may indeed slow the rate of acquisition of the L2 feature, while dissimilar phenomena will increase the rate of acquisition. For example, the English vowel /æ/ was learned at a faster rate than /ɛ/, because the former category does not occur in BP, while the latter is perceptually close to BP /ɛ/. It is also argued that features of the L2 may be acquired at a faster or slower rate by learners of different first language backgrounds

2 See section 3.2 of Chapter Three for a discussion of the universal order of acquisition of phonological elements within the structuralist framework.

depending on the correspondence, or lack of it, between the L1's and the L2 (McLaughlin, 1987).

Avoidance constitutes one of the less obvious manifestations of L1 transfer. Celce-Murcia (1977) in her study of the spontaneous speech of an English-French bilingual child, noted the latter's consistent attempt to avoid physiologically difficult forms, e.g. fricatives. This resulted in the child's use of alternative lexical items in either language which did not contain the difficult sounds. For example, the child produced "couteau" for "knife", and coined a new word "piedball" for "football". What is interesting about this study is that the child replaces a difficult fricative sound /f/ with less difficult stop sounds /p/ and /k/. This confirms Jakobson's (1941/1968) claims that fricatives constitute a more difficult category of sounds and are therefore acquired much later than the less difficult stop category (see section 3.2 below).

Transfer can also manifest itself in the form of hypercorrection which, for example, has been observed among Arabic learners of English (Odlin, 1989). The lack of any contrast between voiced /b/ and voiceless /p/ bilabial stops often results in the substitution of Arabic /b/ for /p/. The attempt to avoid such substitutions leads to "overcompensation" and inappropriate use of /p/ as in "hapit" for "habit", even though /b/ exists in Arabic. L1 influence may also result

in the approximation of certain L2 sounds as already demonstrated by Nemser (1971).

With new insights into the phenomenon of transfer, SLA researchers are not only concerned with the nature of transfer, but also with the conditions which favour or disfavour transfer from the L1 to the L2. Such transferability constraints are examined in the next section.

2.2.3 Transferability of L1 items

There are various views as to what is transferred in the phonic domain. As noted in section 2.2.1. the classical contrastivists consider as transferable phonemes, their variants and the distribution of phonemes and their variants. Generally, the following are considered to be particularly susceptible to transfer: (a) phonetic categories, e.g. vowels, fricatives (Flege, 1987a; Karpf et al., 1980), stops (Abraham and Lister, cited in Mochizuki, 1981; Flege, 1980; Flege and Port, 1981; Gass, 1984) and liquids (Miyawaki et al., 1975; Mochizuki, 1981; Sheldon and Strange, 1982); (b) "low level" phonetic/phonological rules, e.g. phonotactically motivated rules of L1 phonology (Broselow, 1987; Singh and Ford, 1987), allophonic rules (aspiration - Broselow, 1987; Gundel et al., 1986; Karpf et al., 1980), cyclical rules (Rubach, 1984), syllable structure rules (Broselow, 1983, 1987); (c) representations (Ard, 1989); (d) surface phonetic constraints (Eckman, 1981a); (e) Stampean

natural processes (Dziubalska-Kolaczyk, 1987; Major, 1987b), and (f) suprasegmental features (e.g. stress, rhythm and intonation).

In order to specify the conditions which trigger transfer, researchers have invoked various criteria - linguistic, psychological and socio-linguistic. These criteria are discussed below.

2.2.3.1 Linguistic criteria

Attempts to account for transferability on linguistic grounds have focused on structural relationships between L1 and L2, i.e. on their similarities and differences. This practice, as we have already observed, dates back to the heyday of traditional contrastive analysis. While CA may be criticized for its weaknesses, there is no denying that some of its assumptions regarding structural relationships between L1 and L2 have been exploited in more modern SLA research. Recent studies provide evidence to show that L1-L2 differences are a strong factor in interference, even though it has been argued that similarities rather than differences are more likely to cause interference (see section 2.2.1 above).

Altenberg and Vago (1987), for example, studied the speech production of two Hungarian learners of American English. They found transfer errors both at the phonetic and phonological levels. The learners substituted Hungarian sounds for English sounds that do

not occur in Hungarian. For example, they substituted Hungarian [t] for the voiceless interdental fricative [θ] [d] for its voiced counterpart [ð], trilled [r̃] for the approximant [ɹ], high front vowel [i] for mid-high front vowel [ɪ], and mid front vowel [ɛ] for low front vowel [æ]. As it is, Hungarian lacks [θ], [ð], [ɹ], [æ] and [ɪ], but has [t], [d], [r̃], [ɛ] and [i]. It also lacks the schwa [ə] which was produced as [ö] in stressed and poststressed positions, and weakened to [â] in prestressed position.

The learners also applied the Hungarian Voicing Assimilation (VA) rule to English as in:

observed -----> [apsorvɛd]

floods -----> [flʌts]

VA in Hungarian is regressive - the preceding consonant undergoes influence from the following consonant. In the above examples, the obstruents /b/ and /d/ are devoiced to [p] and [t] respectively before the following voiceless obstruent /s/. On the contrary, VA in English operates progressively - the following consonant is the segment which undergoes influence from the preceding consonant. The above examples would be realized as [əbzʒ:vd] and [flʌdz].

In Beebe's (1987a) study of adult Thai learners of English, he found that a range of Thai variants - [r̃]alveolar trill, [ɾ] alveolar flap and [ɺ] lateral flap -were substituted for the English alveolar approximant [ɹ], which does not occur in Thai. Similar

observations have been made for Japanese speakers who were found to produce a range of native language variants - [ɾ], [ɹ] and [ď] - for English /l/ and /r/ (Dickerson, 1976; Dickerson and Dickerson, 1977). The Japanese problem is not only that of differences in the realization of lateral and non lateral sounds, but also differences in the phonemic organization of these sound segments. While /r/ and /l/ occur as separate phonemes in English, there is only one phoneme in Japanese, which is sometimes realized with a lateral-like quality (Sheldon and Strange, 1982). It can therefore be said that while /r/ and /l/ are divergent in English, they converge into one phoneme in Japanese.

Schmidt (1987) studied Egyptian Arabic learners of English and noted that his subjects frequently substituted Egyptian Arabic alveolar fricatives [s] and [z] for English interdental fricatives [θ] and [ð] respectively. Interdental fricatives do not occur in Egyptian Arabic, and as a result the learners substitute sounds in the L1 which are closest to them.

Brière (1968) in an earlier study, examined American English (AE) speakers' production of non English sounds from (a) Arabic (/h/, /ḥ/, /x/, /ɣ/), (b) French (/u/, /ʒ/, /ɛ̃/ and /e/), and (c) Vietnamese (/t'/, /t/, /ɨ/, /ʃ/ and /ɛ/ realized as [ɛ] and [e]) chosen to reflect a specific learning problem for AE speakers. Actually, some of these sounds partially overlap (in terms of shared features) with AE sounds while others do

not. For example, the Arabic laryngeal fricative /h/ is similar to the AE laryngeal fricative [h+], but differs in terms of voicing (breathy voiced), status (phoneme) and distribution (syllable initial). AE [h+] is partially breathy voiced, occurs as an allophone in intervocalic position. French /ɛ̃/ and AE /ɛ̃/ share the features front, unrounded and lower mid, but they differ in that while the former is phonemic and fully nasalized occurring syllable finally, the latter assumes an allophonic status, is partially nasalized, and occurs only before nasals. On the contrary, /e/, /x/, /ɣ/ and /h/ do not have any close associates in AE.

Generally, Brière found that sounds similar to those in the native language - /h/, /t/, /ŋ/, /e/, [ɛ̃] and /ɛ/ - were significantly easier to learn than those which were dissimilar /ḥ/, /ɣ/ and /ï/. A major exception was the voiceless velar fricative /x/ which though non existent in AE was significantly easier to learn than its voiced counterpart /ɣ/. For this segment, the most frequent substitution was L1 [g], along with other variants [?] and [ɸ]. While Brière could not adequately explain the exception, the discrepancy in the results obtained cannot be considered an abnormality if other criteria such as universal markedness are taken into account (see sections 2.4 and 3.6 below).

Flege and Davidian (1984) investigated final stop consonant production in English by speakers of Polish, Spanish and Chinese, and noted that transfer effects due

to differences between native and target languages were more obvious in the Spanish data than in the data for the other groups. While English permits the occurrence of consonant phones in final position, Spanish is very restrictive allowing only laterals, taps, trills, nasals and dental or slightly retroflexed fricatives in this position. The stops /p, t, k, b, g/ do not occur finally in Spanish, and /d/ is ordinarily realized as a voiced or voiceless fricative. On the other hand, Spanish permits /b, d, g/ in intervocalic position but these segments are realized as homorganic fricatives. The Spanish subjects in the study transferred their Spanish Fricativization rule into English and consequently realized word final stops as fricatives more frequently (19% of the time) than subjects in the Polish and Chinese groups who fricativized word final stops less than 2% of the time.

Dreasher and Hsieh (1990) studied Brazilian Portuguese speakers' production of AE sounds which do not occur in the L1 - /θ/, /ð/, /ʃ/ and /æ/. The results revealed that L1 transfer accounted for many of the substitutions that were made for AE sounds. The following Brazilian Portuguese variants - [t], [t̻], [d], [d̻], [s] and [z] - were substituted for /θ/ and /ð/, and /ɛ/ was substituted for /æ/.

Flege and Port (1981) studied the phonetic implementation of the stop voicing contrast in word initial and final positions by Saudi Arabian learners of

English. The results showed that the phonetic implementation of stop voicing in Saudi Arabian Arabic directly influenced the subjects' production of English stop voicing. They produced a duration contrast between the closure intervals of word initial voiced and voiceless stops in English similar to that observed in Arabic, but they did not produce a similar contrast in word final position. This reflects the lack of such a contrast in Arabic.

While the above-mentioned studies emphasize interlingual differences as a determining factor for transfer, there are other studies which demonstrate that similarities rather than differences constitute the major source of interference errors. Major (1987c) observes that transfer is more likely to occur and persist for phenomena which are similar in L1 and L2 than for phenomena which are dissimilar. Flege (1987a) posits that L2 learners produce novel sounds more accurately than those sounds which are perceived to be equivalent in the L1 and L2, because they can avoid using previously established L1 patterns. This position contrasts with that of Brière, cited above.

Flege (1987a) and Flege and Hillenbrand (1987) examined the accuracy with which English speakers produced French /u/, /y/ and /t/ in /tu/ and /ty/ syllables. A contrastive comparison of the sounds in question reveals that French /u/ and English /u/ are similar, differing only acoustically in that the former

is produced with lower frequency (F2) values than the latter. French /y/, on the contrary, constitutes a new phone, having no counterpart in English. French /t/ shares similarities with English /t/, but differs in the sense that it has short-lag VOT (voice onset time) values, while the English /t/ has long VOT values.

The results of the study showed that the new phone /y/ was produced more accurately than /u/. The subjects produced French /y/ with mean F values that did not differ significantly from the values produced by native speakers of French. On the contrary, French /u/ was produced with F values that were significantly higher (i.e. more English-like) than values produced by the native speakers of French. Finally, the subjects' VOT values for French /t/ were longer than values produced by French speakers, corresponding closely to English values.

It is acknowledged, however, that structural properties in themselves do not account for transferability. Therefore, most researchers relate linguistic considerations to universal principles of language acquisition. The concept of markedness, for example, forms the core of Eckman's Markedness Differential Hypothesis (MDH). We shall reserve the discussion of the concept for section 2.4 of our study. But suffice it to say that the consideration of markedness principles alongside structural properties of

the L1 and L2 is thought to provide a better basis for interference predictions (Eckman, 1977).

Gundel and Tarone (1983) predict transfer only in those areas of L1 which accord with the typological characteristics of the learner's interlanguage and "natural" language acquisition processes. Similarly, Zobl (1980) relates the transferability of an L1 item to the interaction between formal properties of L2 and language acquisition principles. An L1 item will only transfer if it conforms more closely to acquisition principles than the corresponding L2 structure. Where the reverse is the case, i.e. where the L2 structure conforms more closely to acquisition processes, then transfer will not occur. However, the transferability of linguistic structures may derive from psychological considerations to which we now turn our attention.

2.2.3.2 Psychological criteria

The psychological approach to transferability emphasizes learner-decision on what should or should not be transferred. Such transferability judgements may be conscious or unconscious (Odlin, 1989). Kellerman (1977, 1983) points out that the learner's perception of L1-L2 typological distance (i.e. the learner's psychotypology) acts to some extent to constrain or to trigger transfer. Crosslinguistic (CL) equivalences and non equivalences form the basis for which transfer may or may not occur (Hammarberg, 1990).

The degree of transfer is assumed to be proportional to the degree of perceived relatedness between the two languages (cf. Weinreich, 1953). Consequently, the transferability of L1 items increases with perceived CL similarities; hence an increase in interference errors, if wrong equivalences are established. Schachter (1983) affirms that interference will occur (a) if the learner wrongly equates a pre-established L1 domain with a current L2 domain, and (b) if the learner chooses the correct domain for both languages, but a hypothesis appropriate for L1 and not for L2.

When learners identify an L2 sound with an L1 sound, they usually substitute the latter for the former - even if both sounds display slight acoustic/ articulatory differences (Flege, 1987a; Flege and Hillenbrand, 1987; Ritchie, 1968). According to Wode (1981) this involves scanning L2 phonological targets for equivalences and non-equivalences, and establishing a crucial measure which separates equivalences from non-equivalences. Only those L2 elements which fall within the "crucial range of equivalence" are matched and substituted by their L1 equivalents. The process by which the learner judges similar phones in L1 and L2 as belonging to the same category is referred to as "equivalence classification".

Equivalence classification is central to Weinreich's (1953) study of languages in contact. He

analyzes interlingual identification in structuralist phonemic terms to account for different types of interference. Identification may be divergent (i.e. a single phoneme category of the L1 is realized as two categories in the L2); convergent (i.e. two categories of the L1 are realized as one in the L2); or simple (i.e. the categories in both L1 and L2 are isomorphic).

Hammarberg (1990) has argued that the similarity scale is gradual, and that the learner may be more or less certain about his equivalent judgements. In other words, there is a dimension of plausibility connected with the equivalence criterion, such that different elements in the L1 may become likely candidates for identification with the L2 target. Hence the substitution of a variety of native language variants for the L2 target. In his study of German beginning learners of Swedish, he found that the long Swedish [u] was identified predominantly with German [ü], whereas the short [ö] displayed greater variation and was identified with [ü] (42%), [u] (27%) and [ö] (11%).

What is interesting about Hammarberg's study is that while the three variants compete for plausibility, the probability of occurrence remains higher for some variants than others. We can therefore argue that there is a hierarchical dimension to the learner's psychotypology: certain elements are ranked higher than others within the "crucial range of equivalence". This assertion is further confirmed by Dreasher and Hsieh's

(1990) study referred to earlier on. Their learners identified English /θ/ mainly with Brazilian Portuguese [t], [d̪], [s], [z] and [f], and English /ð/ with [d̪], [s], [z] and [t]. The results indicate that stops constituted 80% of all substitution errors, while fricatives constituted only 20%. In other words, stops rather than fricatives would be more readily substituted for the L2 segments.

Equivalence classification and its effects have been dealt with in great detail by Flege (and others) in several articles (see for example, Flege, 1987a; Flege and Eefting, 1987: Flege and Hillenbrand, 1987; Flege and Port, 1981; Port and Mitlieb, 1980). The consequences of equivalence classification include (a) the development of inaccurate perceptual targets for L2 sounds with counterparts in the L1, and (b) a resultant "merged system" in which IL productions exhibit values which are intermediate between L1 and L2. As such, L2 sounds with "counterparts" in the L1 are harder to master than those which lack "counterparts". In Flege and Port's (1981) study, for example, Arabic learners of English were shown to produce English stop voicing contrast with VOT values that were shorter than normal English values, but not short enough to be completely Arabic. In other words, such values were neither typically Arabic nor typically English. Caramazza et al. (1973) obtained similar results in their study of French-Canadian bilinguals.

(1990) study referred to earlier on. Their learners identified English /θ/ mainly with Brazilian Portuguese [t], [d̪], [s], [z] and [f], and English /ð/ with [d̪], [s], [z] and [t]. The results indicate that stops constituted 80% of all substitution errors, while fricatives constituted only 20%. In other words, stops rather than fricatives would be more readily substituted for the L2 segments.

Equivalence classification and its effects have been dealt with in great detail by Flege (and others) in several articles (see for example, Flege, 1987a; Flege and Eefting, 1987: Flege and Hillenbrand, 1987; Flege and Port, 1981; Port and Mitlieb, 1980). The consequences of equivalence classification include (a) the development of inaccurate perceptual targets for L2 sounds with counterparts in the L1, and (b) a resultant "merged system" in which IL productions exhibit values which are intermediate between L1 and L2. As such, L2 sounds with "counterparts" in the L1 are harder to master than those which lack "counterparts". In Flege and Port's (1981) study, for example, Arabic learners of English were shown to produce English stop voicing contrast with VOT values that were shorter than normal English values, but not short enough to be completely Arabic. In other words, such values were neither typically Arabic nor typically English. Caramazza et al. (1973) obtained similar results in their study of French-Canadian bilinguals.

learners identify the L2 sounds in terms of these properties rather than others? (Ritchie, 1968).

In as much as conditions which trigger transferability judgements remain unclear (Hammarberg, 1990; Odlin, 1989), the criteria for CL identifications may be broadly categorized as phonological and phonetic (i.e. articulatory, acoustic and auditory) (Flege and Hillenbrand, 1987). James (1986:144) posits that in order to relate the structural characteristics of both languages, the learner must develop a set of associated phonological, phonetic and articulatory properties and a set of rule-governed systematic regularities relating them. An L1 item will only transfer if it meets these constraints.

The learner may categorize L1-L2 sounds in terms of their place and manner of articulation, distinctive features of the Jakobsonian or Chomskyan types, or on the basis of some other set of dimensions and values (Ritchie, 1968). Phonemes, syllables and words constitute phonological units of identification which act as primes of transfer. Brière (1968) has argued that the syllable may be a more central unit of identification than the phoneme or distinctive feature. Selinker (1972) claims that the syllable is a unit of identification at the phonological level.

Kellerman (1983) reintroduces the notion of markedness which rests not on universal linguistic considerations (cf. Eckman, 1977; Greenberg, 1966), but

on learner-perceived markedness of an L1 in relation to a given L2. The emphasis Kellerman places on L1 as the transfer determining factor distinguishes his approach from that of Zobl (1980) which ascribes the strongest impact on transferability to formal properties of the L2 (see p. 35 above). According to Kellerman, an L1 feature perceived as marked (i.e. infrequent, irregular, uncommon) has less transfer potential than an unmarked item. Marked items are considered to be language-specific, therefore non transferable, and unmarked items language-neutral and transferable.

It is worth noting that Kellerman's proposal was not originally intended for phonological analysis, but there is no reason why the concept of learner-perceived markedness cannot be applied to L2 speech analysis. The awareness of language specificity of an item explains why the Dutch speaker would not transfer the voiceless velar fricative [x] and the Anaang Speaker the voiceless labio-velar plosive [k͡p] into English. The specificity or neutrality of the L1 with respect to a particular L2 depends on perceived L1-L2 distance. An L1 item considered to be specific in terms of one L2 may be neutral in terms of another (Kellerman, 1983). Thus, while an Anaang speaker will not transfer [k͡p] into English, he/she will not hesitate to do so when learning Yoruba (which has a similar sound in its phonemic inventory). [k͡p] therefore exhibits language specificity

in relation to English, but language-neutrality in relation to Yoruba.

The learner's psychotypology as well as his/her perception of language specificity and neutrality with respect to a given pair of languages may develop and change with increased contact with the L2, and with increased metalinguistic awareness. Hammarberg (1990) argues that as the learner becomes more acquainted with the L2, the basis for a particular transfer solution may shift, and this will affect the dominant pattern of competing solutions. Different L1-oriented solutions may succeed each other over time, changing along a simplification/elaboration continuum (p.213). What remain remarkably stable irrespective of L2 proficiency are the learner's judgements of the transferability of L1 marked items (Helmond and Vugt, 1984). Certain perceived similarities may remain resistant to eradication, thereby resulting in fossilization (cf. Selinker, 1972) or retardation of the L2 learning process (Zobl, 1980).

2.2.3.3 Sociolinguistic criteria

It has been argued that the social value attached to a native language item also determines its transfer potential. As such, a sound is more or less highly valued according to the setting in which it occurs. The general view holds that less prestigious sounds occur in informal situations, while more prestigious sounds occur in formal situations (Beebe, 1987a). Furthermore, highly valued

sounds are expected to show a higher transfer potential than less valued sounds. In Thai, for instance, /l/, flapped /ɾ/ and trilled /r̆/ appear in informal (e.g. conversation), semi-formal (e.g. careful speech) and very formal (e.g. word list reading) settings respectively. This pattern was demonstrated in Beebe's study of Thai learners of English, which showed a low (3.8%) to a moderate (6.7%) occurrence of flapped /ɾ/ in conversation and listing tasks respectively, but a relatively high occurrence (24.4%) of the trilled /r̆/ in formal listing tasks. This variant did not occur in conversation. /l/ appeared in the listing task, but only at a 6.7 percentage rate.

An interesting dimension to the study was the reflection of native language sociolinguistic patterns in different phonetic environments within a particular task - e.g. word listing. The trilled /r̆/ appeared in initial but not in final position. The latter position showed no native language variants. Instead final "r" was either omitted or correctly produced with 72% accuracy. As demonstrated by Beebe (1987a), initial position "r" variants in Thai have a highly conscious, learned social meaning which final "r" lacks; hence the occurrence of the trilled /r̆/ in initial position and in formal contexts only. Final "r" does not exist in Thai except in loanwords. The social importance attached to initial "r" therefore explains its high transferability rate into English in the said position and context. When informal

and formal tasks were compared, the overall rate of occurrence of /r̂/ was found to be higher than that of /ɾ/ and /l/, thus supporting the hypothesis that highly valued sounds have a greater transfer potential than less valued ones.

Beebe (1988a) points out that these examples represent only a small portion of sociolinguistic transfer. She identifies three main types of transfer: (a) sociolinguistic transfer which involves the transfer of an L1 variable rule as in the above example, (b) transfer of socio-cultural competence, and (c) socially motivated transfer, focusing on the reasons for which transfer occurs. The final type of transfer involves the transfer of a categorical rule in order to fulfil a social psychological purpose in the L2.

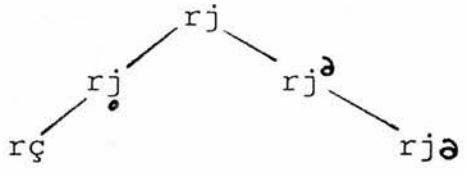
Beebe's study is one of many studies which address the issue of stylistic and contextual variability in interlanguage phonology (see also Dickerson, 1976; Dickerson and Dickerson, 1977; Schmidt, 1987; Tarone, 1983). The main argument is that the learner's IL exhibits systematic variability, which resembles sociolinguistically determined variability in the native language (Dickerson and Dickerson, 1977). We shall examine this phenomenon more closely in section 2.5 of this study. Meanwhile, let us consider the developmental dimension of L2 speech acquisition.

2.3 The developmental dimension

The phenomenon of development in L2 speech acquisition can be seen from different perspectives. Firstly, there has been great interest in exploring sequential development for phonological phenomena as in syntax and morphology (see Dickerson, 1975; Dickerson, 1976; Dickerson and Dickerson , 1977; Hammarberg, 1988b). In the Dickersons' longitudinal studies, it was demonstrated that the pattern of variation observed in the production of target English sounds by Japanese speakers, and its development over time was related to a dimension of increasing phonetic closeness to the target. For example, in the production of /z/ the target variant [z] occurred alongside other variants [s], and zero [0] at Test Time 1 (Dickerson, 1975). At Test Time 2 the proportion of these other variants decreased, and that of the target increased, marked by the appearance of closely related variants [dz] and [z]. Time 3 was characterized by a more marked increase in target production, and a corresponding decrease in close variants. The approximation of the target sounds was also shown to correlate in a systematic way with the nature of the immediate phonetic environment. (See section 2.5 of this study for a fuller discussion of this phenomenon).

While the Dickersons' studies indicate that the acquisitional sequence forms a linear route, Hammarberg (1988b, 1990) has demonstrated that this is not necessarily the case. In his study of final consonant

cluster (/lj/, /rj/ and /rn/) production by German learners of Swedish, the developmental pattern is better interpreted as a set of branches, converging at the target point (insofar as the target point has been reached). The learners, for example, approximate /rj/ from alternative directions, producing the following variants (or "branches") which ordinarily do not form a phonetic continuum: [r], [j], [ri], [rç] and [rjə]. Nevertheless, intermediate values of phonetic approximation to the target can be observed with some of the variants. There is a partially devoiced variant [r̥j] which is intermediate between [rç] and the target, and a variant with a very weak epenthetic vowel [rjᵊ] which is intermediate between [rjə] and the target. The acquisitional sequences for the two "branches" are represented as follows:



(Hammarberg, 1990:210)

The second perspective from which L2 speech development has been considered relates to the reactivation of processes which seem to stem from "universal constraints on human memory/cognition, or from innate characteristics of the human speech production

mechanism or auditory perception" (Flege and Davidian, 1984:324). These processes serve to simplify speech output and make it more manageable for the learner.

Basic to the Creative Construction Hypothesis (CCH), as we have already noted, is the assumption that the same universal cognitive mechanisms operate in L1 and L2 acquisition (Ervin-Tripp, 1974). If L1 and L2 acquisition are governed by the same universal factors, it follows that developmental processes similar to those observed in L1 acquisition will also occur in L2 acquisition as a result of the reactivation of L1 strategies (Alternberg and Vago, 1987; Flege and Davidian, 1984; Johansson, 1973; Macken and Ferguson, 1987; Tarone, 1978; Wode, 1981). Thus, developmental errors made by L2 learners of English will resemble those made by children learning English natively.

Corder (1981) has, however, pointed out that similarities in L1-L2 processes do not necessarily imply similarities in the course or sequence of learning. Others posit that developmental processes are "probabilistic": many, but not all individual learners will show the effect of various developmental processes during speech learning (Flege and Davidian, 1984; Macken and Ferguson, 1987).

It is worth noting that not all L1 processes occur in L2 acquisition. Reduplication (the transformation of model words into structures of identical or repeated CV syllables) and consonant harmony (a process by which two

or more consonants in the same word or syllable are given the same place of articulation) seem to be strictly L1 processes (Macken and Ferguson, 1987). L2 processes similar in part to those in L1 acquisition include terminal devoicing, final consonant deletion, cluster reduction, vowel epenthesis (i.e. vowel insertion in between clusters and vowel addition to final consonants), schwa paragoge, and assimilation processes (Eckman, 1977; Flege and Davidian, 1984; Major, 1987a; Tarone, 1980).

Nonetheless, there are some researchers who claim that L1 and L2 acquisition operate independently in the sense that the processes underlying acquisition are not similar for both languages. Oller (1974) compares consonant cluster reduction processes used in L1 and L2 acquisition, and concludes that while native English learners employ cluster reduction, final consonant deletion and unstressed syllable deletion strategies to simplify syllable structures, L2 learners of English employ epenthesis (vowel insertion) and paragoge strategies.

Ingram (cited in Weinberger, 1987) posits that children learning English natively display deletion strategies before and in lieu of epenthesis. Olmsted (also cited in Weinberger, 1987) claims that children learning their L1 rarely utilize epenthesis as a syllable simplification strategy, but employ deletion extensively. Weinberger (1987) speculates that the reason epenthesis rarely occurs in child language acquisition is that the

time at which *recoverability* is acquired, corresponds to the attainment of adult-like phonetic accuracy. Recoverability functions to prevent excessive ambiguity in language. Epenthesis is said to enhance, while deletion prevents the recoverability of underlying forms of phonological segments.

Tarone (1980) refutes the claim that L2 learners rely solely on epenthesis. In her study, Korean, Cantonese and Portuguese subjects are shown not to rely heavily on epenthesis as a strategy in altering syllable structure. In fact, the learners rely on deletion, epenthesis as well as glottal stop insertion. It must be mentioned though that the Portuguese speakers rely more on epenthesis (80%) than the Korean and Cantonese speakers (30%). Conversely, the latter group uses consonant deletion strategies more often (70%) than the Portuguese group (20%). (This phenomenon is discussed in greater detail below). It is also worth noting that most consonant deletion and epenthesis take place in the production of final consonant rather than of non final consonant clusters.

In Flege and Davidian's (1984) study, Spanish and Chinese subjects were found to delete final stops in English contrary to Oller's claims that L2 learners would favour vowel epenthesis (addition) after such stops. Sato (1984) studied consonant cluster production by Vietnamese learners of English, and found that his

subjects employed cluster reduction and deletion as simplification strategies.

The trend in SLA research has been to study the relationship between developmental and transfer processes (Broselow, 1988; Flege and Davidian, 1984; Hecht and Mulford, 1987; Major, 1986a, 1986b, 1987a, 1987b). This relationship has been interpreted in different ways. While some researchers claim that both processes operate independently, others claim that they interact in interesting ways.

Flege and Davidian (1984; see above) tested two developmental processes (final stop deletion and devoicing) and one transfer process (fricativization) on their Spanish, Polish, Chinese and native English subjects. They hypothesized that (a) Spanish speakers would show a transfer effect by fricativizing stops (especially /b, d, g/) more frequently than the other groups, (b) all three non native groups would show a developmental effect if they devoiced /b, d, g/ more frequently than the native group, (c) Chinese and Spanish speakers would delete stops more frequently than others (developmental effect), (d) Polish speakers would devoice /b, d, g/ significantly more often than others showing a joint effect of L1 and developmental constraints, and (e) Spanish speakers would both devoice and fricativize /b, d, g/, also showing the joint effect of transfer and developmental processes.

All but hypothesis (d) were confirmed. The authors found that the Polish speakers did not devoice /b,d,g/ significantly more often (48%) than the Spanish (43%) and Chinese (30%) speakers. Polish has a phonological rule that devoices final obstruents, while Chinese does not allow some stops in final position. Native Spanish does not have final stops, but a few may be found in loanwords. The fact that the three non native groups devoiced final /b,d,g/ at almost the same rate, irrespective of native language constraints, points to the universal nature of developmental processes. Chinese and Spanish speakers also showed another developmental process: final stop deletion. On the basis of these results, Flege and Davidian argue for the independence of transfer and developmental processes.

However, their Spanish data do show some sort of relationship between the two processes. The Spanish speakers not only fricativize final /b, d, g/ (a transfer process), but they also devoice them to an extent (a developmental process). The two processes seem to be in competition in the realization of these sounds. We are not, however, told which conditions trigger fricativization and which trigger devoicing. The pattern of production may be either sporadic or systematic (see section 2.5 of our study). What is interesting is that while the data may reveal some sort of relationship between the two processes, developmental processes occur

at a significantly higher rate (43%) than transfer processes (19%).

Macken and Ferguson (1987) draw attention to the difficulty in distinguishing transfer from developmental processes for speakers of certain L1's. (See also Flege and Davidian, 1984; Major, 1986a, 1986b). For example, final stop devoicing is considered to be both a transfer process from German as well as a developmental process in L1 English acquisition. Therefore, it is impossible to ascertain whether a German speaker's devoicing of final consonants in English results from transfer or developmental factors. The possibility that one process may be misidentified for another is therefore very high. Any attempt to distinguish transfer from developmental processes in adult L2 acquisition must be preceded by a clear definition of the two (Flege and Davidian, 1984).

Hecht and Mulford (1987) explored the interaction between transfer and developmental processes in the phonology of a six year old Icelandic boy acquiring L2 English. Unlike Flege and Davidian (1984) who do not propose a dual influence of both processes in speech acquisition, Hecht and Mulford actually demonstrate that the two processes may jointly affect a single phone. Their subject not only devoiced final stops (an L1 English process), but also produced them with a strong voiceless aspirated release typical of Icelandic. For example "big" was pronounced as [bik^{h}]. They hypothesize that L2 pronunciations arising jointly from transfer and

developmental processes occur more frequently and persist longer than those arising from just one kind of process. They further propose that the relative roles of both processes will differ depending on which parts of the L2 learners' developing phonology are involved. The following phonological continuum is suggested:

Table 2.1: Hypothesized role of transfer and L1 developmental processes in major classes of phonological segments.

Vowels	Liquids	Stops	Affricates/ fricatives
Transfer processes predominate			L1 developmental processes predominate

According to the continuum, transfer processes predominate for segments to the left, while developmental processes predominate for those to the right. Liquids and stops exhibit a joint effect of both processes, though it is argued that dual influence is stronger for stops than for any other phone. Hecht and Mulford, however, recognize that more research is needed to substantiate this tentative proposal.

Anderson (1983) posits that transfer is most likely to persist if developmental processes would predict the same result. Thus, terminal devoicing would persist for German learners of English. We can also assume that the Polish speakers in Flege and Davidian's (1984) study referred to above, would find it difficult

to overcome terminal devoicing, at least much more than their Spanish and Chinese counterparts.

The influence of the L1 and the mechanism of transfer vary in the course of L2 speech acquisition (Kellerman, 1983; Leather and James, 1991). This implies that the impact of the L1 is stronger at certain stages of learning than at others. Major (1987a) posits that at the early stages of acquisition transfer processes predominate at the expense of developmental processes. As acquisition proceeds interference processes give way to developmental processes which increase and then decrease over time. Major notes that the cross-over point or rate of change will vary from individual to individual, and that some learners will not go through all the stages of acquisition.

With regard to the stylistic domain, Major claims that interference processes decrease while developmental processes increase and then decrease as tasks become more formal (i.e. from text to sentence to word). He also claims that transfer is more likely to occur and persist for phenomena which are similar in L1 and L2, and developmental processes for dissimilar phenomena in the two languages (Major, 1986a, 1987a, 1987b). The learner tends to be more aware of such differences and therefore consciously attempts to overcome L1 interference. The relationship between transfer and developmental processes

form the locus of the Ontogeny model which is examined in Chapter Three.

According to Wode (1981), however, those L2 elements which do not meet specifiable similarity requirements, and are therefore not substituted by L1 elements will be acquired via developmental processes. His study of German learners of English showed another type of developmental process which did not involve deletions or neutralization of an L2 contrast. He observed that German children (aged 4-9 years) realized English /r/ as a [w]-like phone instead of the uvular [ʁ] of the L1. This process known as gliding (i.e. the replacement of liquids (/r,l/) with glides (/w,j/)) is well documented in the L1 acquisition of English (see Locke, 1983). Generally, [w] replaces both /r/ and /l/, while [j] replaces /l/.

The reactivation of this strategy has not been reported for adult L2 learners of English. Beebe (1987a) however, observed that Thai speakers produced [wɹ] for initial /r/ 3% of the time. He describes this variant as an "r" with a "w" colouring, and considers it to be a new variant - i.e. an approximation and therefore neither L1 nor L2-like. He also observed that [w] was produced for final /r/. While children acquiring English natively are reported to replace /r/ with [w], it has not been shown that they do so in final position. Thus, it is possible that the child L2 learners in Wode's study may have been reproducing [w] for /r/ substitutions heard in the speech

of native English mates (Flege and Davidian, 1984). Adult learners often prefer one of the liquids as Locke (1983) indicates. It may be argued therefore, that what takes place in child L2 acquisition does not necessarily typify adult L2 learning.

However, a study by Piper (1984) seems to indicate that child L2 and adult L2 acquisition involve the same processes, at least as far as the production of stops and fricatives is concerned. Piper investigated the phonological development of 4 and 5 year old children of different L1 backgrounds - Portuguese, French, Cantonese, Punjabi, Italian, Vietnamese and Hungarian - learning English as their L2. Like adult learners, the children made use of substitution (e.g of interdental fricatives /ð/ and /θ/ with [d] and [d̪]), and final consonant devoicing and deletion strategies. Piper stresses the evolving nature of these processes. In other words, they are not static, but change in character and pattern over time, resulting in the gradual approximation of the target segment (cf. Dickerson and Dickerson studies cited above). For example, the following developmental sequence was produced: [d] ---> [d̪] ---> [ð]. The children made use of all the strategies irrespective of L1 background, a fact which again points to the universal nature of the processes.

Nevertheless, despite their "universal basis" (cf. Flege and Davidian's 1984 study cited above), the same developmental processes **will not always** occur in the

speech of L2 learners of different native language backgrounds. That is, native language constraints may, to some extent, determine which developmental processes occur in the L2 speech, or at least what proportion of these processes are involved.

In Tarone's (1980) study discussed above, it was observed that the critical variable in the choice of strategies by Korean, Cantonese and Portuguese learners of English was the L1 background. While the Cantonese and Korean speakers preferred consonant deletion as a strategy, the Portuguese speakers favoured epenthesis. According to Tarone's descriptions of syllable structures in the three languages, Cantonese is more or less an open-syllable language, with most syllables consisting of a consonant-vowel (CV) structure. "Close" syllables which may end in one of a limited number of consonants may occur. Clusters occur across word boundaries but not within syllables. Portuguese is, relatively, an open-syllable language as well, allowing few consonants in final position. Clusters occur within syllables, and in slow speech an [i] paragoge is inserted between clusters. In native acronyms and loanwords with final consonants, a vowel is added to the consonant (Major, 1986a). Korean on the other hand is primarily an open-syllable language with a more complex syllable structure.

From the above analysis it is predictable that Cantonese and Korean speakers would most likely use deletion, and Portuguese speakers epenthesis as

strategies for simplifying syllable structures. In a related study, Anderson (1987) found that his Vietnamese subjects used mainly deletion strategies to simplify syllables, and that this could be attributed to native language constraints. Eckman (1981a) reports epenthesis as the main strategy employed by his Mandarin learners of English.

Another developmental process - final consonant devoicing - was predicted on the basis of L1 constraints, and confirmed for Cantonese speakers in a study by Edge (1991). According to Edge's analysis, Cantonese allows obstruents (stops) in final position, but they are invariably unreleased, and a glottal stop is usually inserted if the following word or syllable begins with a vowel (p. 379). The data were coded as target and non target (devoicing, deletion, epenthesis, fricativization, glottal stop insertion and others). The overall results showed that devoicing accounted for 67% of non target variants produced by the Koreans. Glottal stop insertion and deletion occurred at an almost equal rate, while very few cases of epenthesis were observed. The study also included Japanese speakers. However, their data, contrary to Eckman's (1981b) claims, did not confirm the hypothesis that Japanese speakers would favour epenthesis over other developmental strategies, as a result of native language surface phonetic constraints.

As it is, the results provided for Cantonese speakers by both Tarone and Edge are conflicting. This

discrepancy may be traced to two major factors. Firstly, both researchers used different tasks to elicit data. Tarone sampled elicited natural speech using just one task in which the subjects were required to describe orally a sequence of pictures, narrating a story in the process. Edge, on the other hand, employed an instrument that sampled various types of speech, using various tasks which differed in levels of formality. The tasks ranged from high formality (list reading) to low formality (riddle discussion).

Firstly, by using a wider range of tasks, Edge was able to elicit as many learner strategies as possible. Certain tasks have been found to favour the occurrence of some strategies over others. For example, devoicing and epenthesis are more likely to occur in formal speech, and deletion in informal (i.e. connected) speech (Edge, 1991; Weinberger, 1987). This probably explains why Tarone did not report any instances of devoicing, given that she elicited informal data. Edge's study confirmed that deletion was an important factor in connected speech for the Cantonese speakers.

Secondly, the focus of each study was different. The data reported were restricted to data necessary to support conflicting hypotheses about IL rules employed by Cantonese speakers. Edge's study was a partial replication and extension of Eckman's (1981b) which reported evidence of IL rule of schwa paragoge for Japanese speakers, and one of terminal devoicing for

Cantonese speakers. Tarone, in the attempt to identify some of the processes that shape the IL syllable structure, tested Oller's (1974) claims that L2 learners would favour epenthesis as a strategy. She may, therefore, not have tested for devoicing. Moreso, in addition to transfer, she proposed the universal CV syllable structure as an explanation for the use of different strategies by her subjects. This then raises the question as to how far the L2 speech is constrained by the general phonetic and phonological properties of sounds and sound systems. The role of language universals in L2 acquisition forms the focus of the next section.

2.4 Language universal constraints

Some of the factors which influence L2 speech acquisition have been attributed to the inherent general properties of the item to be learned, or to universal constraints on language (Hammarberg, 1988b; Leather and James, 1991). In discussing language universal constraints on L2 acquisition, we shall consider the universal CV pattern and sonority relations in syllable structures, and the markedness of phonetic/phonological forms.

That the simple open CV syllable may be a universal articulatory/perceptual unit has been demonstrated by Tarone (1980) whose study we have discussed at length. While the majority of errors identified by Tarone were attributed to language

transfer, there were a few which could not be traced to L1 interference, and which resulted in a modification towards the basic "unmarked" CV syllable. On the basis of these findings, Tarone concluded that the preference for the open (CV) syllable seemed to operate as a process independent of language transfer.

Tarone's study was replicated by Hodne (1985) who investigated English final consonant and cluster production by native speakers of Polish. The tasks included an oral interview and a narration based on a videotape which was shown to the subjects. Hodne, like Tarone, attributed a majority of the errors in her study to language transfer, and also found a number of errors which supported Tarone's "Open-Syllable Hypothesis". Nevertheless, the results of this study are less categorical than Tarone's, with about 50% of the overall modifications showing a clear movement towards the CV syllable pattern. Hodne, therefore, stresses that the universal CV rule is applied with variable consistency by learners of different language backgrounds.

However, while the preference for the open syllable operates as a variable rule in IL phonology, with the L1 background influencing the extent to which it is applied, Hodne suggests that the patterns of modification found in the data

> might best be analyzed not in terms of processes unique to IL phonology, but rather in terms of universal patterns of glottalization and epenthesis found in the speech of persons experiencing stress, whether they are speaking in their native language or in a second language (Hodne, 1985: 405; see also Labov, 1977).

Other studies conducted along Tarone's lines include Sato (1984) and Benson (1988). Nonetheless, these studies provide little evidence to support the claim that the CV rule operates independently of language transfer. Sato's Vietnamese subjects consistently deleted final consonant clusters which could be the result of L1 transfer or a universal preference for the CV syllable. Benson (1988) also studied the production of English final consonants by Vietnamese speakers. The results of his study suggest that in addition to L1 influence, the preceding vocalic context plays a significant role in determining any universal CV preference in acquisition.

Benson's study differs from the three studies cited above in that it includes native language data, gathered and analyzed under the same experimental conditions as the IL data. L1 data gathered in this manner seem to provide a more "authentic" basis with which to compare performance, than extanct descriptions provided by linguists. On the other hand, this kind of native language data may not reveal "everything" about the structures being investigated due to time constraints, and the type of native data elicited.

Broselow (1983) presents evidence to show that the different ways in which Egyptian and Iraqi Arabic speakers resolve initial consonant cluster problems in English is a direct reflection of the L1 syllabification rules, and to some extent universal constraints. The Egyptian speakers, for example, pronounce "floor" as [filor], while the Iraqis pronounce it as [iflor]. The L1 syllabification rules are such that the Egyptians insert a vowel to the right of an unsyllabifiable consonant, creating an open syllable, and the Iraqis insert the vowel to the left of the consonant creating a closed syllable.

However, the Egyptians deviate from this pattern in one class of words, those beginning with **s-stop** clusters (e.g study), which are produced with the epenthetic vowel to the left, i.e. before "s": [istadi]. Broselow (1988) traces this phenomenon to the fact that **s-stop** clusters are the only clusters which violate the universal Sonority Principle in English syllable onsets (see section 3.6.2. below). As a result, they are treated as single units at some level of representation, and are not capable of being divided by epenthesis. Learners, constrained by this sonority principle are forced to analyze all clusters that violate the sonority hierarchy as complex.

The role of the sonority hierarchy in shaping syllable structure in learner speech has been emphasized

by Tropf (1987) in his study of the acquisition of sy llable final consonants, initial and final clusters by Spanish learners of German. The learners resolved cluster difficulties by using epenthesis and reduction strategies. The general tendency was to drop the less sonorous consonant and retain the most sonorous one. For example, initial /ts/ and /pf/ clusters were reduced to /s/ and /f/ respectively or to some other fricative. The same principle applied in final consonant production in that the less sonorant segments tended to be omitted more frequently than the more sonorous ones. Plosives were most frequently omitted, followed by the fricatives, and then nasals and liquids.

Interestingly, plosives and fricatives constitute a natural class of obstruents, while nasals and liquids constitute a natural class of sonorants in the sense of Chomsky and Halle (1968). Tropf posits that differences between the two classes are greater than differences within the classes. The difference in omissions between the lateral and nasals was less obvious than that between plosives and fricatives. On the whole, Tropf's subjects omitted plosives more frequently than fricatives, and nasals more frequently than the lateral.

It is worth noting that the sonority value of a particular segment interacts with L1 factors in the sense that it is only operative where the L1 does not possess equivalent cluster types (Leather and James, 1991). Counter evidence for this claim would be provided if the

L1 and L2 have similar cluster types, and in spite of this the L2 learner engages in cluster reduction. Evidence for sonority as an independent operative factor is provided in cases where the L2 violates the sonority hierarchy. In German this applies to final plosive-fricative clusters, e.g. [klaps] (Tropf, 1987). Here too the pattern is to omit the less sonorant plosive and retain the more sonorant fricative.

The concept of sonority has also been exploited by Broselow and Finer (1991). The aim of their study was to test whether clusters closer in sonority are harder to learn than those which are wider apart. An element of implicational universal markedness, defined in terms of a set of parameters and parameter settings given by Universal Grammar (UG), was introduced (see section 3.3). Clusters whose members are closer in sonority were considered to be more marked than those whose members are wider apart.

The subjects consisted of Korean and Japanese learners of English who were required to produce sentences containing initial *pr, br, fr* and *py, by, fy* clusters. According to the hierarchy of sonority relations (see section 3.6.2.), **consonant + r** clusters are more marked than **consonant + y** clusters. Thus, while Japanese and Korean have the unmarked parameter setting, English has both the marked and unmarked setting. **Cr** were predicted to cause more difficulty than **Cy** for Japanese and Korean speakers. In general, the results

confirmed the predictions, though a breakdown of the results revealed a relative degree of difficulty among the clusters. In other words, the clusters did not exhibit the same degree of difficulty. For example, learners performed relatively better with *py, by* and *pr* clusters than with *br* clusters, though the number of errors corresponded to the degree of markedness of the clusters based on sonority relations. Most errors were committed for *br* clusters because this is the most marked combination of the four. The more highly marked *fy* and *fr* clusters exhibited a higher rate of errors.

An interesting dimension to the results is the fact that initial *b*, but not initial *p* occurs in Japanese, which should make *br* easier than *pr*. This is not the case as we have seen. The results suggest that universal constraints (i.e. markedness) can offset the effects of the L1. Evidence from the study further suggests that learners have neither retreated to the unmarked setting for a given parameter nor have transferred, wholesale, the L1 setting. Instead, their responses reflect a parameter intermediate in markedness between those of the L1 and L2. This effect occurs both when the L2 employs a less marked setting than the L2, and when the L2 is more marked than the L1.

Another study which investigates the role of universal constraints on L2 speech acquisition is Dreasher and Anderson-Hsieh (1990) (see section 2.3). The aims of their study were (a) to determine whether

language universals could predict the order of difficulty of American English sounds - /θ/, /ð/, /ŋ/ and /æ/ - that do not occur in Portuguese, either phonetically or phonemically, and (b) to investigate the relative difficulty of these sounds for native speakers of Brazilian Portuguese. The results revealed the following hierarchy of difficulty which confirmed the predictions that were made:

Table 2.2: Hierarchy of difficulty for American English sounds.

(a) /n/ and /ŋ/	easiest
(b) /θ/	intermediate
(c) /ð/ and /æ/	most difficult

Less marked sounds were produced more accurately than more marked sounds. Native language sounds were substituted for target sounds as a way of resolving the difficulty. Interestingly, the substituted L1 variants are shown to be universally less marked than the target segments. For example, less marked stops and fricatives were substituted for more marked interdental fricatives (cf. Jakobson, 1941/1968; Menyuk, 1968). As already noted, stops occurred at a higher frequency than fricatives. Stops are said to be less marked than fricatives. Thus, wherever substitutions occurred, it was the less marked L1 variants which were produced.

The relationship between typological markedness and crosslingual differences has been dealt with by Eckman (1977). His Markedness Differential Hypothesis (MDH) predicts difficulty where the target language differs and is typologically more marked than the native language, and no difficulty where the reverse is the case. Most research investigating the interplay of L1 and Universal factors in L2 speech learning have focused mainly on the syllable structure.

Anderson (1987) tested syllable structure difficulty for Chinese and Arabic learners of English. He predicted that the relative degree of difficulty of English consonants for each group would correspond to the length of the cluster: the longer the cluster, the more marked and therefore the more difficult it would be for the learners. He also predicted that final clusters would be more difficult than initial clusters. Final position is considered to be universally more marked than initial position (Hammarberg, 1988a). The predictions were supported as both the Chinese and Egyptian groups had more difficulty with longer "marked" clusters. Final clusters were also more difficult than initial clusters. Given that the Egyptian Arabic syllable structure is closer to that of English than is the Chinese syllable structure, Anderson also predicted that Arabic speakers would perform better than the Chinese speakers on English Syllables that resembled those of Arabic. This prediction was also confirmed.

In Benson's (1988) study referred to earlier on, more marked English clusters were shown to cause difficulty for L1 Vietnamese speakers whose language allows only the less marked singleton consonants (cf. Greenberg, 1965). What is more, as in Broselow and Finer's (1991) study, the cluster types did not have the same level of difficulty. Relative difficulty was predicted on the basis of the relative markedness of the clusters. Benson therefore concluded that in order to predict learning problems in the L2, the markedness relations between L1 and L2 must be considered alongside markedness relations within the L2 itself for a given phenomenon.

However, the predictions of markedness have not been borne out in other studies. Hammarberg (1988a) in a study of German learners of Swedish found no cross-language markedness correspondences which could account for the difficulty of certain Swedish items. For example, the German palatal non sibilant fricative [ç] is no less marked than the Swedish palatal sibilant [ɕ] which ranked among the most difficult segment types.

In the same vein Altenberg and Vago (1987) could not attribute the difficulty their Hungarian subjects had with English Voicing Assimilation to the markedness differential between the two languages. In other words, Hungarian Regressive Voicing Assimilation is no less marked than English Progressive Voicing Assimilation.

Eckman's MDH is discussed in greater detail in Chapter Three of this study.

In this section we have examined the role of universal constraints on L2 speech acquisition, and the way in which these constraints interact with the L1 and developmental factors. In the next section we turn our attention to the phenomenon of variability in IL phonology. We consider the extent to which the factors discussed here account for such variability.

2.5 Variability in interlanguage phonology

Central to Selinker's (1972) Interlanguage Hypothesis is the concept of variability. When acquiring an L2, the learner internalizes a system of variable rules, some of which apply on certain occasions and some on other occasions. In Dickerson's (1975) study of adult Japanese learners of English, the target sound /z/ was produced correctly as [z] when preceded by a vowel, but was replaced by a series of variants - [s], [dz] including [z] when followed by silence. The phonetic environment constituted a strong determining factor on the variable production of /z/. Flege and Davidian's (1984) Spanish subjects fricativized English /b, d, g/ in final position on a number of occasions and also devoiced them on a few occasions (see section 2.2 above). Though it is not indicated which occasions triggered fricativization and which triggered devoicing, we can assume that the preceding vocalic environment may have

been a factor. Sources of variability are discussed below.

Two types of variability have been identified in IL phonology: situational and contextual variability. Variants of a target sound are likely to occur in different phonetic environments and to be responsive to different styles. Studies on variability were pioneered by Labov (1970) who dealt mainly with variability in the first language. He examined the speech patterns of New Yorkers and observed that the frequency of certain socially marked sounds (e.g /θ/) varied along a continuum of styles - (a) casual speech, (b) careful speech, (c) reading texts and minimal pairs - and that performance differences were relative to the degree of attention paid to speech. Casual speech demands less attention to speech, while minimal pairs demand a maximum degree of attention. Fewer instances of target /θ/ were observed in the former than in the latter.

The same correlations hold in non-native (i.e. L2) speech. Tarone (1979, 1982, 1983) has argued for style shifting in L2 speech acquisition, determined by the degree of attention paid to language form, which in turn depends on the type of activity (e.g. various elicitation tasks - imitation, translation, word listing, text/dialogue reading and oral interviews).

In Dickerson's (1975) study referred to above, data were elicited from Japanese speakers on three separate occasions over a nine month period using

different tasks: (a) free speech, (b) dialogues and (c) word lists. The range of variants produced for /z/, including the correct variant, depended on the degree of formality of the tasks involved. The target variant or other linguistically close variants occurred most frequently in (c), the most formal task, and least frequently in (a) the least formal task. The frequency of occurrence of target or closely related variants was intermediate in (b). These results reveal that the target and closely related variants appeared in situations which allowed for audio-monitoring, and linguistically distant variants in situations where audio-monitoring was absent.

Developmentally, there was an increase in the proportion of target and target-like variants. There were more occurrences of the target forms at Time 3 than at Time 1. The same findings were replicated in a similar study by Dickerson and Dickerson (1977) which investigated the production of English /r/ by Japanese speakers. The target sound was produced 50% correctly in free speech, and 100% correctly in word listing.

Gatbonton (1978) studied the variable production of interdental fricatives /θ/, /ð/ and the velar glide /h/ by French Canadian learners of English, and found that the variants produced for these segments correlated well with the particular speech style involved. The more formal the style (e.g minimal pair reading as opposed to spontaneous speech), the greater the number of target-like realizations of the segments in question. The

following variants including the target were produced for (a) /θ/ ---> [θ], [tθ] and [t], (b) /ð/ ---> [ð], [dð], [d] and a zero [0], (c) /h/ ---> [h], [?] and [ɸ].

The significance of task type in style shifting was also demonstrated by Schmidt (1987). He reported that Egyptian Arabic learners of English exhibited more accurate use of interdental fricatives [θ] and [ð] in formal than informal tasks. In the latter, less target-like sibilant sounds [s] and [z] were used. These results re-echo those he obtained in an earlier study of Egyptian learners of English (see Schmidt, 1974).

Sato (1984, 1985) has demonstrated that task-oriented style shifting is not unidirectional in that formal styles will not always exhibit the most target-like production. His longitudinal studies of word final consonant and consonant cluster production by Vietnamese learners of English reveal that style shifting can proceed in any direction. For instance, while more target-like realizations of final consonant clusters occurred in the careful rather than in the vernacular style, the opposite was the case for final consonant production. There was more target-like production in the vernacular style than in the careful style. Developmentally, task variation was more prevalent for clusters than for final consonants.

The question remains as to what part L1 transfer plays in such variable production. We may also ask what roles developmental and universal processes play in such

variable production. Concerning developmental processes (see section 2.3), Major (1986a, 1987a) has pointed out that as style becomes increasingly formal interference decreases and developmental factors increase and then decrease, to give way to target production. For example, in final consonant cluster production, text reading yielded more transfer than developmental errors, while the word list and normal sentence reading yielded only developmental errors. This pattern was observed among beginning as well as advanced Portuguese learners of English, though there were significant differences in the proportion of transfer and developmental errors between the groups. The interaction between transfer and developmental processes was dealt with extensively in section 2.3, and therefore need not be repeated. From the universalist perspective, fortition and lenition processes are said to be stylistically conditioned (see section 3.4 below). Fortition processes (e.g insertion) are more common in formal styles than are lenition processes (e.g. reductions and deletions) which are more common in casual speech.

However, variability in L2 speech production is generally investigated from the perspective of transfer. In Dickerson and Dickerson (1977) the target sound was frequently replaced by native Japanese variants [ɾ̆], [ɹ] and [ɾ] particularly in free speech. This supports the claim that interference is more likely to occur in casual speech than in formal speech styles.

Major (1987b) maintains that the pattern of variability in IL production may reflect an on-going process in L1. He reports that speakers of Brazilian Portuguese generally pronounce English "meadow" with a final [o] in word listing but may pronounce it as [u] as style becomes more casual due to the process /o/ --> [u] / __#, which occurs in running speech in Brazilian Portuguese. This process is not observed in formal word listing.

Generally, casual speech seems to reflect more L1 interference than careful speech, because the IL is said to become more permeable to the L2 system in increasingly formal situations (Tarone, 1979). However, L1 interference may persist in formal styles where socially prestigious variants used in the L1 are transferred to equivalent settings in the L2. In Schmidt's (1974, 1987) studies mentioned above, the variable production of interdental and alveolar fricatives were shown to reflect style shifting in Arabic: [s] and [z] occur in colloquial Arabic (associated with informal situations), while the interdentals [θ] and [ð] occur in classical Arabic (associated with formal situations). The phenomenon of transfer in style shifting has also been dealt with by Beebe (1987a). Beebe's study was discussed at length in section 2.2.3 of our study.

As with situational variability, contextual variability constitutes a continuum: some linguistic contexts yield certain variants while other contexts yield other variants (Ellis, 1985b). Phonological

segments change from one phonetic shape into another by virtue of their position in a word or syllable, or under the influence of surrounding sounds. Contextual variability in L1 acquisition was examined by Labov (1970). SLA researchers (as cited above) have also examined this phenomenon.

In Dickerson (1976) it was shown that the learners produced different variants for /l/: [ɾ], [r̃], [ɫ̆] including the target variant [l], depending on whether /l/ preceded or followed high, mid or low vowels, preceded or followed consonants. There were nine phonetic environments, In environment A, for example, which consisted of /l/ + a low vowel /a/, the subjects attained categorical (i.e. 100%) use of /l/ from test time to test time. In environment B which consisted of /l/ + a mid vowel /e/, two variants [ɹ] and [l] were produced, the two occurring in free variation at Times 1 and 2. The target variant occurred in free variation with the Japanese variant. There was, however, an increase (80%) in target production at Time 2. Time 3 showed 100% target production, with the native language variant disappearing completely from the learners IL.

The same findings were replicated for /r/ production in Dickerson and Dickerson (1977). Target production was more accurate (75%) when /r/ preceded low vowels, and less accurate (50%) when /r/ preceded mid vowels. There were no target variants before high vowels. Instead, more native-like variants appeared in

this position, and less before mid vowels. There were no native variants when /r/ preceded low vowels. It can be seen from the studies examined above that in many cases the learners' output contain phones which are those in the native language. As Dickerson and Dickerson (1977) point out, these phones usually appear in the same environments as in the L1.

In Gatbonton's (1978) study variants of /θ/, /ð/ and /h/ were conditioned by the segments which preceded or followed them. For example, target /ð/ occurred more frequently after a vowel than after the four consonantal environments. Gatbonton has argued that this may be due to the fact that the vocalic environment contrasts highly with /ð/ than the consonantal environments. In the latter contexts of /ð/ it seems to be the case that those sharing the same features as /ð/ (i.e. [+voice]) yield more target-like production than those which do not (i.e [-voice]) (Gatbonton, 1978).

Other studies which examine phonetically conditioned production of L2 sounds include Gillette (1980), Henly and Sheldon (1986), and Sheldon and Strange (1982). Interestingly, these studies examine variability in production as well as in perception. Both the production and perception of L2 speech segments are phonetically conditioned, but in different directions as we shall see in the next section.

2.6 Production and perception of L2 speech

The study of L2 speech acquisition is not restricted to production but includes perception. As in L2 speech production, the main focus has been on the influence of the native language on the perception of L2 phonological categories. SLA researchers have also been interested in the relationship between production and perception. In other words, to what extent can the L1 account for perceptual inaccuracies in L2 speech acquisition? Are production and perception isomorphic, or are they independent one of the other?

It has been suggested that adults, like children have an innate ability to discriminate between human speech sounds, but that they tend to perceive the segmentals and suprasegmentals of the L2 in terms of the categories of their L1 (Bailey and Haggard, 1973; Best and Strange, 1992; Brière, 1968; Broselow, Hurtig and Ringen, 1987; Flege, 1987a, 1987b; Gass, 1984; Gillette, 1980; Kohler, 1981; Lado, 1957; Mackain et al., 1981; Major, 1987b; Mochizuki, 1981; Sheldon and Strange, 1982; Strange and Jenkins, 1978).

Non-native phones are perceived in terms of their similarities and dissimilarities to the native phones. The categorization of target language stimuli parallels the native language categorization in so far as the native language categories have phonetically and (phonologically) similar counterparts (Scholes, 1968). Conversely, the categorization of the L2 will not

parallel that of the L1 where differences exist between the two languages. The phonological and phonetic inventories of languages differ. Two phones which occur in the L1 may be completely absent in the L2. In another instance, both the L1 and L2 may share phones which are phonologically contrastive in the former, but occur in contextual or free variation in the latter, rather than being used to distinguish meaning (Mackain et al., 1981). In the reverse situation, the two phones may occur as contextual or free variants in the L1, but as phonological contrasts in the L2. The question remains as to whether speakers of languages which do not contrast these phones phonologically are able to perceive and produce them distinctly when learning a language that makes such a contrast.

Japanese, Korean and Cantonese monolingual speakers are known to have difficulty distinguishing English liquids - /r/ and /l/ - which do not occur contrastively in their native languages (Borden et al., 1983; Henly and Sheldon, 1986; Logan et al., 1991; Mackain et al., 1981; Miyawaki et al., 1975; Mochizuki, 1981; Sheldon and Strange, 1982; Zimmerman et al. 1984). Miyawaki et al. (1975) tested the perception of /r/-/l/ contrasts by Japanese and American English speakers using synthesized speech stimuli. The results showed that while the Americans discriminated the contrasts with high accuracy as expected, the Japanese speakers performed only slightly better than chance. However, when the

stimuli were presented in isolation, no difference was found between the two groups.

MacKain et al. (1981) have demonstrated that native-like perception of /r/-/l/ contrasts is possible with highly proficient L2 learners. They found that conversationally proficient adult Japanese speakers were able to identify and discriminate the contrasts like the native English speakers, whereas their conversationally inexperienced counterparts were not able to do so.

English speakers are reported to have difficulty with non-native contrasts between Czech retroflex and palatal fricatives (Trehub, cited in Leather and James, 1991), Thai voiced and voiceless unaspirated (Lisker and Abramson, also cited in Leather and James, 1991). Boatman (1990) reports difficulty in the perception of non-native vocalic /u/-/y/ contrast by American English learners of French. Bohn and Flege (1990) found that German native speakers learning English had difficulty distinguishing the non-native /ɛ/-/æ/ contrast.

The argument is therefore that the lack of experience with a given phonological contrast results in poorly defined perceptual boundaries separating the two members of that contrast (Mackain et al., 1981). The poor perception of L2 boundaries often results from the absence of such boundaries in the L1. Hence the assimilation of target language categories into native language categories.

Best and Strange (1992) tested the identification and discrimination of three synthetic series of American English approximants contrasts - /w-j/, /w-r/ and /r-l/ - by Japanese and American speakers. Based on a comparison of Japanese and English phonology with respect to the segments in focus, they predicted (a) that the Japanese /w-j/ boundary would be shifted toward /j/ relative to the American boundary because Japanese /w/ is unrounded and is more similar to English /j/, and (b) that the /w-r/ series would tend to be labelled as /w/ rather than as /r/, i.e. /r/ would be assimilated unequally to a single Japanese phoneme category. No clear predictions were made about the /r-l/ boundary, though it was expected that learners would assimilate the contrast to a single category. It was also predicted that highly proficient Japanese speakers would identify all three contrasts in a pattern similar to that of American speakers, while the low proficiency group would identify the contrasts at a much lower rate. These predictions were all confirmed.

Assimilations are thought to be made on the basis of perceived gestural similarities or acoustic-phonetic similarities (Brown and Goldstein, 1986; Flege, 1987a, 1987b), though the distinction between gestural and acoustic-phonetic properties are difficult to make given that both are confounded in the signal (Best and Strange, 1992). However, it is argued that if a novel phonetic contrast is to be learnt by adults, selective attention

must be allocated to new acoustic-phonetic dimensions that were previously unattended to in the L1 (Logan et al., 1991). Therefore, Japanese speakers must learn to allocate their attention to acoustic cues that differentiate /r/ and /l/ in English (Logan et al., 1991).

While differences in L1-L2 phonological/phonetic constraints may pose production as well as perception problems, it is not quite clear what relationship holds between the two phenomena. Does the inability (or ability) to perceive new contrasts translate directly into the inability (or ability) to produce such contrasts? In other words, does the learner with better perception of the target language perform better at production tasks than a learner who inaccurately perceives the target language? (Major, 1987b).

The early contrastivists (Carroll, 1968; Dalbor, 1959/60; Lado, 1957) argued for an isomorphous relationship between perception and production in that learners were assumed to produce only those sounds they could perceive. Thus, any observable production difficulties reflected perception difficulties as well. Nemser (1971), whose study of Hungarian learners of English was discussed at length in section 2.2.1 above demonstrated that this is not necessarily the case.

Recent views do not therefore recognize a one-to-one relationship between production and perception in L2 phonological acquisition. One view holds that L2

perceptual mastery precedes production mastery in learning new phones and contrasts (Borden et al., 1983; Major, 1987b; Neufeld, 1980, 1988). In this case the learner's mental representation of sounds, is identical with the target language. Neufeld (1980, 1988) studied the L2 speech of English Canadian learners of French, and found asymmetry between phonological perception and articulatory production. Performance on the former was much better than performance on the latter. Neufeld's studies suggest that adult learners can acquire a native-like or near native-like perception. An accent would therefore be the result of production difficulties (see also Major, 1987b). He further adds that the reason why L2 learners retain a foreign accent in production is because they are exposed to bad L2 models.

Borden et al.'s (1983) study of the production and perception of /r/-/l/ contrasts by Korean learners of English also points to a perception lead over production. The study which included training sessions showed no improvement in production, but showed significant improvement in perception, particularly in identification rather than in discriminatory functions. Perception scores were higher than production scores, a fact which suggests that training in perception did not carry over into production. On the contrary, in an earlier study, Pimsleur (1963) found that discrimination training not only improved discrimination between L2 minimal phonemic pairs, but also that this gain carried over into

production. Evidence of positive transfer from perception to production is also provided by Mueller and Niedzielski (cited in Leather and James, 1991).

A different view claims the reverse, i.e. that production is well ahead of perception (Brière, 1968; Gass, 1984; Goto, 1971; Schneiderman et al., 1988: Sheldon and Strange, 1982). Brière (1968) (see section 2.2.3.1 above) has argued that perceptual mastery is not necessarily a causative factor in the acquisition of productive skills. He found, for example, that while his American subjects were able to produce Arabic pharyngeal /ḥ/ and laryngeal /h/ accurately, they could not perceive the contrast distinctly. Goto's (1971) investigation of Japanese speakers' perception and production of English /r/-/l/ contrasts showed that the learners' production ability far exceeded their auditory perception ability, even for their own utterances. L2 proficiency did not seem to affect the pattern of performance, though the highly proficient learners performed better than the less proficient learners.

Goto's study was replicated by Sheldon and Strange (1982), with the same results, though the subjects in the latter study generally performed better than those in the former. In a reanalysis of data in the Borden et al.'s (1983) study, Sheldon (1985) observed that perception and production abilities interrelate in different ways, and that the relationship seemed to depend on the length of

time spent in an English speaking country, in this case, the United States.

Another study which demonstrates a production lead over perception is Gass' (1984) study of Farsi, Italian, Portuguese, French, Korean, Japanese and Thai learners of English. She found that while they could not perceive /p/-/b/ boundaries, their production of the contrast approximated native speaker norms. The study by Bohn and Flege (1990) referred to above seems to indicate that the interaction between perception and production can go either way depending on the proficiency level of the L2 learners. Their experienced German learners of English produced the /ɛ/-/æ/ contrast better than they did perceive it. On the contrary, the inexperienced Germans failed to produce the contrast, but identified the contrast accurately. Bohn and Flege have suggested that "perception is more resistant to L2 experience than production because speech production is more subject to social control than speech perception" (1990:52).

The multiple views expressed above indicate that researchers do not agree on the nature of the relationship between perception and production in adult L2 acquisition. Therefore, inferences about L2 production cannot be drawn from L2 perception and vice-versa. However, there is some consensus about the nature of the perception-production relationship in child language acquisition whether first or second. For L1 acquisition, it is generally argued that perception

precedes production (Eimas, 1975; Menyuk, 1977; Smith, 1973: Strange and Broen, 1981; Strange and Jenkins, 1987). Eimas (1975) and Strange and Broen (1981) showed that young children can perceptually differentiate /r/ and /l/ prior to producing the contrast. Menyuk (1977) has argued that the acquisition of perception occurs gradually. Williams (cited in Borden and Harris, 1980) studied L2 acquisition of English by Puerto-Rican speaking children and noted a shift in the production of both English and Spanish words, accompanied by a shift in perception. Production and perception shifted towards the target language.

It has been shown that the perception and production of non-native contrasts vary as a function of the word position in which the contrasts occur (Gillette, 1980; Henly and Sheldon, 1986; Sheldon and Strange, 1982). Goto's (1971) study, referred to above, examined the production and perception of /l/ and /r/ in different word positions (initially in singleton and consonant clusters, and medially), but failed to take into account the variable performance that resulted from the two sounds occurring in different phonetic environments. Not all positions were equally difficult for the Japanese speakers. His claim that Japanese speakers cannot identify or produce /r/ and /l/ contrasts is therefore true for /r/ and /l/ in some positions, but not in others (Mochizuki, 1981).

Of the environments tested in Gillette's (1980) study of Japanese and Korean speakers, the medial and final environments were found to be generally easy, while the initial and consonant cluster environments were difficult for both language groups. There was an increase in accuracy in all environments from test time to test time for both groups except the Koreans' identification of the contrast in initial position. In the consonant cluster position, *br* clusters were more difficult than *bl* clusters. Various vowel environments also appeared to affect the accuracy of identification of /l/ or /r/. High vowels were more difficult than low vowels.

For perception the general trend is that performance is poorest in initial position (for singleton consonants and consonant clusters) as well as in intervocalic position. On the contrary, the final position (singleton consonants and clusters) yield better performance (Logan et al., 1991). It must be emphasized that this pattern is extremely general since some studies do reveal minor deviations. For example, in Logan et al's study of the perception of /r/-/l/ contrasts by Japanese speakers, the intervocalic environment was one of the easiest environments, apart from the final environment. In Mochizuki's (1981) study the identification rates by Japanese speakers of /r/-/l/ words produced by an American English speaker revealed the following order of increased accuracy: intervocalic, cluster, initial and final. In the study by Sheldon and Strange (1982) the

order from most difficult to least difficult was CC > intervocalic > initial > final position. /l/ was misidentified more than /r/ in initial position, and /r/ more than /l/ in CC position.

For production the order of difficulty as reported by Mochuziki (1981), ranked from easiest to hardest is : initial /r/-/l/, final CC /l/, intervocalic /l/, initial CC /l/ and final CC /r/. In Sheldon and Strange (1982) CC's are more difficult in perception than in production. The asymmetry between perception and production is further highlighted by these results.

The question remains as to what extent phonetically conditioned production and perception of target sounds can be predicted on the basis of native language constraints. In Japanese, for example, /r/ occurs initially and medially, but not finally, and CC's are not allowed (Sheldon and Strange, 1982). Japanese has no /l/ phoneme though [ɺ] occurs as a flapped allophone of /r/, which is a flapped sound itself. A simple contrastive analysis would predict difficulty in word final position, and ease in initial and medial positions, but this does not seem to be the case for Japanese speakers.

In an attempt to explain these phenomena, Dissosway-Huff (cited in Henly and Sheldon, 1986) proposes the Duration Hypothesis which claims that the duration of English /l/ is longer in final than in initial position. For this reason, the Japanese can

perceive English final /l/ more accurately than he/she does initial /l/. This explanation cannot therefore be accounted for on the basis of L1 experience, as it would imply that speakers of other languages (e.g. Korean) who have difficulty perceiving the /r/ and /l/ distinction should have no difficulty with the final position.

Henly and Sheldon (1986) report a different pattern: their Cantonese subjects identified /r/ and /l/ best in initial and medial positions. Initial CC and final positions were more difficult. Final /l/ caused the most difficulty because the learners employed the Cantonese /l/ perceptual template for judging all English /l/ variants. Cantonese /l/ is similar to the non velarized variant of English /l/ (i.e. word / syllable initial /l/). In English final velarized [l] differs from initial [l] in that the F1 and F2 values are more characteristically close for final [l] and wide apart for initial [l] (Henly and Sheldon (1986). In the study the liquids that matched the Cantonese perceptual template were identified as /l/, while those that did not were identified as /r/. The latter included final position /l/. Reliance on the Cantonese /l/ template facilitated the perception of English /l/ in non final position, but not in final position.

What these findings show is that native language phonological systems affect L2 acquisition to varying degrees, directly or indirectly. For Japanese and Cantonese speakers, the perception of /r/-/l/ contrasts

pose difficulty because both sounds do not exist as abstract phonemic categories in their L1's. While Cantonese has /l/ in its phonemic inventory, Japanese has the /r/ phoneme, which in turn has variants including the lateral flap [ɺ]. However, it seems to be the case that fine phonological details in Cantonese play a more obvious role in the perception of English /r/-/l/. The effect of L1 acts to filter out acoustic characteristics of English /r/ and /l/. The effect of filtering depends on the existence and/or distribution of /r/ or /l/-like phonemic categories in the learner's native language. According to Logan et al. (1991) the consistent finding that the phonetic context affects the perception of /r/ and /l/ indicates that listeners may not necessarily proceed directly from the phonemic categories of their native language to those of the target language, but may rely on intermediate, context-sensitive phonetic categories when initially learning a new contrast (p. 881).

However, fossilization may occur in the course of acquisition. It is to this phenomenon that we now turn our attention.

2.7 The fossilization of L2 phonology

One of the phenomena which characterizes L2 speech acquisition and distinguishes it from L1 acquisition is fossilization. According to Selinker (1972:215)

> fossilizable linguistic phenomena are items, rules and subsystems which speakers of a particular NL will tend to keep in their IL relative to a particular TL, no matter what age of the learner or the amount of explanation and instruction he receives in the TL.

An example of fossilization is the maintenance of the French uvular /r/ in English by native speakers of French. Learners who fossilize are considered not to have reached the end of the IL continuum (see section 3.7 below).

Selinker (1972) views fossilization as a mechanism which exists in what he calls the "latent psychological structure", and is latent in the brain (p.212). This latent structure is activated when the learner tries to produce an L2 utterance. Fossilized structures may be realized as correct or incorrect L2 forms. The former occur when the leaner's IL development regarding a particular feature assumes the same form as in the L2. If the IL does not assume the same form as the target, then fossilized errors will occur. Fossilizable structures may re-emerge in the productive performance of an IL when seemingly eradicated. These structures appear in the IL when the learner is tackling a new and difficult aspect of language, or when he is in a state of anxiety (Selinker, 1972).

One of the central issues in IL phonology therefore is that of fossilization in the IL of adult L2 learners. This does not rule out fossilization in the IL

of child L2 learners as well, for "under certain conditions the child's progress in a second language may be susceptible to the effects of fossilization as an adult's" (Selinker and Lamendella, 1978:148).

Two questions to which researchers have addressed themselves relate to the inevitability of fossilization in L2 learning and the causes of such fossilization (Tarone, 1978). Researchers are divided in their opinions. Scovel (1969) maintains that no adult ever achieves perfect pronunciation in the L2. Therefore, permanent fossilization in the adult learner is inevitable. His arguments derive from the basic assumptions underlying Lenneberg's Critical Period Hypothesis (CPH) as discussed below. Adjemian (1976) also argues for the inevitability of fossilization in adult IL, resulting from a permeable IL system which allows the influx of foreign rules, or the overgeneralization or distortion of an IL rule.

Opposing views hold that phonological fossilization is by no means inevitable. Hill (cited in Tarone, 1978) reports of South American and New Guinea Indians who learned several L2's as adults and achieved native-like proficiency in them. Neufeld (1988) has pointed out that L2 learners can be taught to acquire native or near native proficiency in L2.

Several explanations have been offered to account for the causes of fossilization in IL phonology. Tarone (1978) classifies these proposals into three broad

groups: physiological, psychological and socio-emotional. Perhaps the most important physiological explanation for fossilization is the CPH. It is claimed that language acquisition is adversely affected after brain lateralization which occurs at the onset of puberty. With the completion of cerebral dominance, the brain is said to lose its capacity for language learning, and this loss seems to affect phonological acquisition more than the acquisition of syntax or lexis (see section 2.2). Krashen (1973) has argued that lateralization takes place long before the onset of puberty, i.e. at the age of five. However, available evidence from SLA research does not seem to support the claims of CPH.

A psychological explanation of IL fossilization centres on Piaget's stage of formal operations during which adults begin to consciously construct knowledge about the world. Thus, instead of simply "acquiring" L2, they tend to "learn" L2, abstracting rules of grammar and consciously applying them.

The psychological Habit-formation theory claims that fossilization occurs as a result of negative transfer of L1 habits into the L2. For some reason it seems that negative transfer operates to make IL phonology resistant to change. This argument has limitations since it is fairly obvious that transfer constitutes only one of the processes which shape IL phonology, and can be positive as well.

According to socio-economic explanations, the point at which the learner's IL systems fossilize is directly controlled by the cessation of the learner's acculturation into the target society (Schumann, cited in Selinker and Lamendella, 1978). Thus, fossilization occurs because of the learner's lack of empathy with the native speakers and culture of the L2. Conversely, fossilization will not occur as long as the learner continues to identify psychologically and socially with the native speakers and culture of the target language. Guiora et al. (1975) claim that pronunciation is a much more sensitive indicator of empathy than syntax or morphology. They also claim that because children have more fluid ego boundaries, they are more likely to identify with speakers of the L2 than are adults who have rigid ego boundaries. The problem with this argument is that empathy does not lend itself to any form of empirical measurement. With so many proposals being made to explain the phenomenon of fossilization, the causes of fossilization still remain unclear and hard to pin down. Moreso, fossilized language structures are very difficult to identify.

2.8 Summary and conclusion

This chapter has examined research trends in the area of interlanguage phonology. A review of the literature has shown that the L2 learner's phonological system is a complex interaction of factors. While first

language transfer remains the predominant factor, it interacts constantly with other factors such as developmental processes and language universal constraints. The complex nature of transfer itself has been demonstrated. The behaviourist notion of the phenomenon has been abandoned for a more cognitive oriented concept which recognizes the active participation of the learner in the transfer process.

The interaction of processes produces a variable interlanguage system. The literature on variability has been reviewed with emphasis on linguistic as well as stylistic variability. The concept of variability has been extended to the perception of phonological elements. The main focus, however, is on the relationship between perception and production. While different proposals have been made the general consensus is that production and perception are not isomorphous. Lastly, we have considered the phenomenon of fossilization. The question as to the inevitability of fossilization has been addressed. On the whole, the arguments remain inconclusive given that fossilizable elements are difficult to identify.

CHAPTER THREE: THEORETICAL ISSUES AND FRAMEWORKS

3.1 Introduction

While the preceding chapter has been mainly concerned with the review of studies conducted in L2 phonology, the phenomena considered have been discussed in relation to particular theoretical concepts such as habit-formation, hypothesis-testing, sonority, equivalence classification, markedness and variability. In the present chapter we reconsider some of these proposals relating them to specific theories of phonology and phonological development. As such this chapter provides a detailed theoretical background to the issues examined in Chapter Two.

It is claimed that the characterization of the process of acquisition is impossible without a coherent theory of language, and in this case a theory of phonology (Chomsky, 1981; Ferguson and Garnica, 1975). A phonological theory should allow for a comprehensive formal description of language (both universal and specific), which should be consistent with psychological reality (Kolaczyk, 1987). In other words, the theory should not only provide an explicit descriptive framework for the formulation of phonological descriptions, but should also be psychologically adequate (Chomsky, 1965, 1981).

A phonological theory is considered to be psychologically adequate if it provides a mentally plausible account of L1 acquisition, which also serves towards the explanation of L2 acquisition. Data from the L2 can in turn serve to validate or modify the theory at various levels of phonological analysis (James, 1989). In essence, the relationship between phonological theory and L2 research is a symbiotic and mutually beneficial one. The problem is that no theory of phonology has yet met the criteria for descriptive or explanatory adequacy as spelt out by Chomsky (1965, 1981).

In the following sections, we shall consider some theories and models that have been proposed, their principal concepts as well as the limitations. We shall begin by examining the structuralist theory which posits the existence of a universal order of acquisition and emphasizes the implicational relations that hold therein. This will be followed by a discussion of the generative and natural phonology theories both which provide a more mentalistic view of phonology and its development.

The extent to which these theories account for L2 acquisition is evident in the separate models that have been developed to account specifically for L2 data. These include the Ontogeny model, the Markedness Differential Hypothesis and the Sonority Hierarchy. They are examined in subsequent sections. The Labovian model will be considered in the final section. James (1989) notes that this model is essentially atheoretical in its orientation

in that it makes no phonological claims about its findings. It therefore fails to fulfil the requirement that a model of phonological development be related to a phonological theory. Nonetheless, the model is interesting in that unlike the others which deal mainly with the psycholinguistic aspects of acquisition, it emphasizes the sociolinguistic dimensions.

Each of these theories/models constitutes a potential framework for the present study. However, given that no theory or model can account for all aspects of phonological development, we propose to adopt an eclectic approach by selecting those aspects which account most adequately for the phenomena that we intend to investigate.

3.2 Structuralist theory

Perhaps, the best known and most influential theory of phonological development was formulated by Jakobson (1968). Working within the structuralist framework, Jakobson attempted to explain the acquisition of phonology on the basis of linguistic universals, structural laws which underlie every modification of language, individual or social (Ferguson and Garnica, 1975). These laws "determine the inventory of phonemic systems and the relative frequency, combinatorial distribution and assimilatory power of particular phonemes" (Macken and Ferguson, 1987:5).

Within Jakobson's framework, the phonological system consists of a universal hierarchy of distinctive features arranged in a strict pattern of successive dichotomous branchings or oppositions. These features are not absolute, but relative properties of phonemes. It is worth noting that Jakobson does not conceive of the phoneme as a mental entity, but as a linguistic abstraction, a theoretical construct on the phonological level (Hyman, 1975). Phonemes and their features are represented in a matrix, with features for the rows and phonemes for the columns (see Jakobson 1968).

The relationship within the stratified phonological system is such that the existence of an entity (or contrast) in a lower layer implies the existence of the entity (or contrast directly) above it. This hierarchy determines the phonemic systems of the languages of the world. Thus, the presence of a particular contrast in a language implies the presence of all those contrasts ordered above it.

Jakobson claims that the development of child phonology is governed by the same general structural laws called "laws of irreversible solidarity", which determine the universal and invariable hierarchy of layers in phonemic systems. He therefore proposes an invariant order of acquisition which cuts across all languages. Children acquiring language, no matter what language, will follow the same general order of acquisition of features and oppositions. However, while the order

remains invariant, the rate of development may vary between children.

A part of the hierarchy predicts, for example, that children will not acquire the phonemic contrast between velo-palatal and labial/dental consonants (/k:p/ or /k:t/) before contrasting dental and labial consonants (/b:d/ or /p:t/). In other words, they will not acquire a contrast which ranks low on the hierarchy before the one(s) ordered above it. The laws of implication referred to above therefore apply.

Phonological development begins with the labial stage, usually the optimal labial stop /p/, which occurs simultaneously with the optimal wide vowel /a/. The contrast between the two units in succession establishes the universal CV syllable (Ferguson and Garnica, 1975). The labial stage is followed by the opposition nasal/oral /p:m/ and then the opposition labial/dental /p:t/.

Generally, stops are acquired before nasals, followed by fricatives, liquids and glides. Affricates are also acquired late. In the early stages of acquisition stops are substituted for fricatives. When fricatives do appear in the child's repertoire, /s/ is said to precede all others. Jakobson makes no firm predictions about the sequence of acquisition of liquids (/l/ or /r/), and glides (/j/ or /w/). However, it seems generally accepted that in English /w/ is acquired first and /r/ last, and that in the transition period, different segments may be substituted for one another,

e.g. /w/, /l/, /j/ for /r/ and /j/ for /l/ (Hawkins, 1984: 281). Front consonants are said to be acquired before back ones (e.g the contrast /p:t/ before the contrast /p:k/), and usually serve as substitutes for back consonants. Voiceless consonants precede their voiced counterparts, while back rounded vowels precede front rounded vowels.

The general hypothesis is that segments that are comparatively rare and physiologically difficult are acquired last, while those that are common and physiologically less difficult are acquired first. (Jakobson, 1968:57). The order of acquisition is therefore tied to the frequency distribution of segments among the world's languages. The frequency of such segments is, in turn, determined by the nature of articulatory demands that are made on the speech organs. These facts have implications for the markedness concept in that phonological elements are acquired in an increasing order of markedness. Unmarked segments often occur early in language acquisition even when no models for them exist in the adult system, while marked segments occur late (Jakobson, 1968; Kiparsky, 1972).

Jakobson maintains that the acquisition of phonological systems operate as mirror images of each other in that the last segments to be acquired are the first to be lost in language decay. For example, liquid segments which emerge late in acquisition are the first to be affected in the decay of phonological systems,

either as a result of linguistic change through simplification, or of the speech pathology of individuals (Brakel, 1983).

In essence, phonological development from Jakobson's perspective is an additive process, involving not just the successive acquisition of phonological oppositions which proceeds "through the universal feature hierarchy from the most general contrast to the finest and rarest contrast" (Macken, 1980:145), but also the elaboration of the feature hierarchy marked by increases in (a) the number of distinctive features in a phoneme and the number of phonemes in the system; (b) the maximum number of phonemes in a word; (c) the number of possibilities for the distribution of phones; and (d) the maximum number of phonemic distinctions within a word (Ferguson and Garnica, 1975:164). Jakobson (1968) points out that the child's perception of the adult system is incomplete at the onset of speech: the complete system is acquired through the gradual unfolding of phonological oppositions over time.

Jakobson's theory has been criticized on various grounds. We shall consider only those criticisms which are relevant to our study. Ferguson and Farwell (1975), for example, question the notion of the acquisition of segments. They argue that early acquisition of phonology is not segment-based but word-based (i.e holistic); the features that are acquired are allocated to the lexical item and not to the segment. This, in their view

explains why features of the respective target items do not necessarily appear in the same order as in the target. For example, in the attempted rendition of the word "pen" by a child "K", the nasal feature which occurs finally is produced randomly in all positions. Ferguson and Farwell argue that the child's difficulty is not in identifying the phonetic/phonological features of the target, but in producing them in the proper sequence.

Whether initial phonological acquisition is segmental or holistic - and in spite of Ferguson and Farwell's disclaimer, there is evidence to show that it is segmental (Smith, 1973) - the point remains that segment-based coding of features develops at some stage in the acquisition process (Wode, 1988). If this were not so, then first language segments would not transfer into the L2. German speakers are noted for devoicing final (but not initial or medial) voiced stops and fricatives in English. These speakers devoice the same segment in the L2 which is devoiced in the L1.

With regard to the proposed order of acquisition, the facts appear to be mixed. While there is proof of an invariant sequence, individual variations among learners, with respect to certain oppositions, do not seem to justify its universality. Though stops are generally acquired before fricatives, it is not true that the sequence for stops -/p/ > /p:m/ > /p:t/ > /p:k/ - will be followed invariably by all children. The opposition dental/labial may precede rather than follow the

opposition oral/nasal for some children (Leopold, 1947 cited in Macken, 1980). The dental consonant /t/ may precede the optimal labial consonant /p/ in early acquisition, while the opposition dental/velar may occur earlier than expected (Menn, 1976 cited in Macken, 1980). Oppositions may merge, such as in the merging of /b/ with /d/. In this case, only the dental member of the pair occurs, and serves as a substitute for the labial member (Menn, 1976).

Similarly, although fricatives are frequently replaced by stops in early acquisition, some children do not follow this pattern. Many rarely substitute stops for fricatives other than the interdentals, /θ/ and /ð/ which are acquired quite late (Leopold, 1947). In English [d] or [t] is a common substitute for [ð]. Comparatively, other fricatives rather than stops are likely to replace difficult fricatives (Ferguson, 1977). Cross-linguistically, the patterns differ. For example, while English children replace /f/ with /s/, Spanish children replace /f/ with /p/. The explanation for these differences lies in the phonological systems of Spanish and English. Contrary to Jakobson's claims /s/ is not always acquired first. Rather, /f/ occurs first in the speech of some children. The general order of acquisition for fricatives is as follows: /f, s, ʃ, v, z, θ, ð, ʒ,/ (Ferguson, 1977:265).

Nevertheless, the fact that the details of segmental acquisition differ in some respects from those

predicted by Jakobson does not necessarily invalidate the universal order. The general sequence holds for most children. Stops are acquired before fricatives, fricatives before affricates, front consonants before back consonants, back rounded vowels before front rounded vowels, and voiceless consonants before their voiced counterparts. Furthermore, the relationship between order of acquisition and the notion of markedness holds.

Jakobson has also been criticized for omitting important aspects of phonological development such as the acquisition of oppositions in environments other than initial, the acquisition of clusters and the relationship between production and perception (Macken, 1980). Given that different positions will trigger different patterns of acquisition, the order proposed by Jakobson may not apply across environments. Moreso, segments are not acquired the same way in clusters as they are as single segments (Ferguson, 1977). For example, /s/ may be readily acquired in initial position as a single segment, but omitted when it forms a cluster with another consonant as in [mouk] for "smoke". /s/ may also occur in one type of cluster (e.g s+l), but not in another (e.g s+m). With regard to the relationship between speech production and perception, different views hold, though it is generally agreed that in child language acquisition perception precedes production (see section 2.6 above).

The question now remains as to what extent Jakobson's theory applies to L2 phonological development. We shall consider two main claims of the theory, which are that (a) there is a universal order of acquisition, and (b) phonological elements are acquired in an order of increasing markedness, such that the acquisition of marked elements imply that of unmarked ones but not vice-versa. The two claims are intertwined in the sense that the universal order of acquisition reflects the increasing markedness of phonological elements.

While there are relatively few longitudinal L2 studies, available evidence seems to indicate that the Jakobsonian order of acquisition does apply to some extent in L2 acquisition. Generally, those segments that emerge late in L1 acquisition, and are less frequent (i.e less basic) in the world's languages are difficult to acquire in the L2. Interdental fricatives and affricates, for example, often cause difficulty for L2 learners. Hence, the substitution of other variants in their place. As we saw in section 2.2.3.1 of Chapter Two, earlier acquired stops and fricatives were often substituted for interdental fricatives, with stops predominating over fricatives. Johansson (1973) concludes from a study of the L2 acquisition of Swedish by speakers of different L1's that certain sounds are more difficult than others:

> ...the same vowels which appear as phonemes in children's speech and which are most basic in the languages of the world, are also reproduced with fewest phonetic deviations (1973: 159).

The more basic back rounded vowel ([ɤ]) is produced with fewer phonetic deviations than the less basic front rounded vowel ([ɸ]), the former being substituted for the latter, as Jakobson predicted. Liquid segments which emerge late in L1 acquisition are also particularly difficult for some L2 learners (see Dickerson and Dickerson, 1977).

However, in examining the order of acquisition in L2 phonology, we must take into account what the learner has already acquired in the L1, as this is likely to be transferred into the L2. A French speaker learning English has no difficulty with the voiceless palato-alveolar fricative /ʃ/ because it occurs in his/her L1 (Major, 1987b). Universally, this is one of the less frequent segments, and it emerges late in L1 English acquisition. However, as a result of the native language influence, the French speaker is able to produce /ʃ/ earlier than expected and with no difficulty. For the Arabic speaker whose L1 has no voiceless bilabial plosive /p/, the acquisition of this segment in English constitutes a problem. While /p/ emerges quite early in L1 English, it occurs late in the IL English of the Arabic speaker, often after /b, t, d, k, g/ have been acquired (Major, 1987b). These examples indicate that the order of acquisition in the L2 may vary slightly from the universal order manifested in children, as a result of the learner's native language.

The conclusion to be drawn from the above facts, on the one hand, is that as with L1 acquisition, the applicability of Jakobson's theory to L2 phonological development has some limitations. The effect of the L1 must be taken into account. On the other hand, there is no doubt that the implicational markedness relations which hold between phonological elements constitute an important factor in the explanation of L2 acquisition (cf. Eckman, 1977). Phonological markedness is discussed as a theory in its own right in section 3.6 below.

To summarise, this section has examined phonological acquisition from the structuralist perspective. The locus of the theory is that the development of phonological systems is governed by the same universal structural laws which determine the inventory of phonemic systems. While this claim may be applied (with some limitations) to both first and second language acquisition, it is worth noting that, on the whole, the theory deals only with the surface constraints of language. The theory of generative phonology which is discussed in the next section proposes a more abstract view of language and phonological development.

3.3 The generative approach

The theory of generative phonology (GP) aims to describe the ideal native speaker's knowledge of the sound system of his/her language. Several versions of the

theory have been postulated. In the Standard (i.e. SPE) version this implies knowledge of a transformational rule system which maps Underlying Representations (UR's - i.e. the systematic phonemic level) onto surface Phonetic Representations (PR's - i.e. the systematic phonetic level) (Chomsky, 1965, 1966; Chomsky and Halle, 1968). Phonological representations are said to be mentally constructed by the speaker-hearer and underlie actual performance in speaking and understanding.

Underlying representations (UR) consist of surface syntactic information and lexical representations of the morphemes, words, phrases and sentences. Phonetic representations consist of linearly ordered phonetic segments whose surface structure is derived from the UR through rules. These rules are said to operate upon segments in strings of morphemes in utterances. They are ordered and apply linearly, though some may apply cyclically - i.e. repeated in cycles. The theory of cyclical phonology has been elaborated upon by Rubach (1984).

The concept of two levels of sound structure is borrowed from the late 19th century linguist, Bauduoin de Courtenay (1895), who distinguished between phonemic structure (a representation of the sound properties of an utterance which differentiate it from others), and a phonetic structure (the actual vocalization of the phoneme) (Anderson, 1979). The central innovation of classical GP is the considerable attempt at making

explicit the principles relating phonological to phonetic representation. The theory provides a detailed description of the nature and process of transformations. URs are considered to be "deeper" and more abstract than originally conceived by De Courtenay. The degree of abstractness is demonstrated in the claim that phonological deep structure contains forms which are deleted from surface representation, e.g. the word **king** is said to have the underlying structure /kiŋg/ from which /g/ is deleted, giving [kiŋ] in the surface structure. It is also claimed that the same form underlies words with a common morpheme, e.g. **divine** and **divinity**. The abstract form for these words is said to be /divain/.

The URs and PRs are given in terms of feature matrices specifying the grammaticality of the vocal apparatus involved in the articulation of the corresponding sound (Kenstowicz and Kisseberth, 1979). The task of phonological rules is therefore to assign features from one level to the other. Chomsky and Halle (1968) stress that at the PR level features are no longer binary and distinctive, but may be multivalued and non distinctive. They, however, remain binary and distinctive at the UR level. Distinctive features are not predictable by phonological rules, while redundant features are. The concept of features is a legacy from Jakobson which has been incorporated into the generative phonology framework with some modification. Thus, while

Jakobson describes features on the basis of the acoustic characteristics of sounds, Chomsky and Halle describe them on the basis of articulatory characteristics.

In summary, classical GP theory is concerned with explicating the relationship between URs and PRs, by examining the nature of formal phonological rules, the ways in which they interact and the distance between representations in phonology. This version of GP has had some flaws. Transformations appear to be too powerful. As such there are no restrictions on the number of operations that they can perform. The outcome in many cases are predictions of processes which are not attested to in any phonological system (Kaye, 1989). Moreso, there are no apparent connections between URs and PRs, and the notion of deep structure is questioned (James, 1980; Kohler, 1971: Wardhaugh, 1970). Criticisms of classical GP theory has led to its being revised and extended in several versions.

The central concept of current generative theory is the Universal Grammar (UG), "the system of principles, conditions, and rules that are elements or properties of all human languages" (Chomsky, 1976:29). In the case of phonology these are "principles governing possible sound systems for human languages, the elements of which they are constituted, the manner of their combination and the modifications that they undergo in various contexts" (Chomsky, 1988:26). The theory claims that UG is inherent in the human mind, being made up of substantive and

formal universals. The former consist of fixed features such as distinctive phonological features, and the latter of abstract principles governing possible rules and parameters of human language. UG theory is therefore concerned with the speaker-hearer's knowledge of principles which apply to all languages, and parameters which vary within clearly defined limits from language to language (Cook, 1985).

However, the theory of principles and parameters has not been developed for phonology to the degree of explanatory potential that has occurred in syntax (James, 1990). Most applications of UG principles to the sound system seem to be restricted to non-linear phonology of which there are several versions (e.g. autosegmental - Goldsmith, 1976, 1990; CV-Tier - Clements and Keyser, 1983; Kahn, 1980; metrical - Liberman, 1979; see also James, 1987, 1989). Common to these theories is the claim that phonological representations consist of independent levels (or tiers) of features (syllabic, stress, tonal, segmental) linked to each other by association lines which are subject to well-formedness conditions (Katamba, 1989).

Some of the principles and concepts identified in the syntactic domain have, however, been extended to the phonological domain. The notions of government and constituent structure, for example, have been explored by Kaye et al. (1990) in their theory of syllable structure. Government is defined as "a binary, asymmetric

relation holding....between two skeletal positions" (p.198). Certain segments have governing properties and are associated to governing skeletal positions. Others are governable and are associated to skeletal positions that are governees. Syllables are broken into syllabic constituents which are governing domains in themselves.

Kaye et al. propose the existence of three syllabic constituents which include the Onset (O), the Nucleus (N) and the Rhyme (R), the Nucleus being governed by the Rhyme. Not any two segments may co-occur within a constituent as their permitted position is determined by their governing properties. The authors maintain the universal nature of these syllabic constituents. Languages may, however, vary along parameters of branching or non branching Rhyme, Nucleus or Onset. Possible combinations of syllables seem to be determined by the relations holding between constituents. For example, the exclusion of branching Rhymes in a particular language inventory means the exclusion of branching Nuclei and Onsets. As such, only five settings of parameters describe the syllabic inventories of the world's languages (Kaye, 1989).

The notions of government and constituent structure may be linked to the principles of Structure-Dependency, Projection and Empty-Category which have been borrowed from syntax. The Projection Principle, for example, require that phonological units and their feature specifications be fed via the lexical level into

higher phonological levels. Kaye et al. (1990) assert that "governing relations are defined at the level of lexical representation and remain constant throughout phonological derivation" (p. 221). Within the metrical phonology framework stress patterns are said to be projected via the lexical level into higher prosodic levels (see Selkirk, 1984). Borrowing from this framework James (1987, 1989) in his version of non-linear generative phonology posits that units and their features are inserted via "lexical composition conditions" into prosodic and rhythmic hierarchies. He, however, stresses the need for an indepth exploration of the Projection and Structure Dependency Principles in order to establish their relevance to sound structure (James, 1990).

One of the few principles proposed specifically for phonology has come from autosegmental phonology (James, 1990). Known as the Obligatory Contour Principle (OCP), it states that adjacent (i.e. consecutive) identical autosegments are prohibited, unless they are separated by word boundary (Goldsmith, 1990:23). Initially applied to tones and intonation this principle has been extended to segmental analysis, the details of which we shall not delve into (but see Goldsmith, 1976, 1990). Another phonological principle is the Sonority Principle, the discussion of which is reserved for section 3.6.2. However, recall that this principle was applied to studies by Broselow (1983), Broselow and Finer (1991) and Tropf (1987) reviewed in Chapter Two.

UG theory has different implications for phonological developmental in that we are no longer concerned with the acquisition of transformational rule systems. Language acquisition entails applying UG principles to a particular language and setting the values that are appropriate for each parameter (Cook, 1985). With exposure to a particular L1, the child fixes parameters for the said language, from the universal set provided by UG.

Parameters that are set with the aid of the UG form the core grammar of the language, and are acquired first. However, not all parameters are core: every language contains elements that are not constrained by UG. These constitute the peripheral parameters of the language, and are acquired last because they require more input. In other words, "a core language is a system derived by fixing values for parameters of UG, and the periphery is whatever is added on in the system actually represented in the mind of the speaker-hearer" (Chomsky, 1986: 147).

Related to the core-periphery dichotomy is the notion of markedness. Core grammar is considered to be marked relative to peripheral grammar. Markedness may also be seen in terms of the choice between settings of parameters. As such a parameter may have marked and unmarked settings (e.g. marked and unmarked settings for sonority sequencing or syllable types).

Language acquisition may therefore be described as involving an initial "state of mind", **So**, which consists of principles and parameters of UG, and a final "state of mind", **Ss**, consisting of UG instantiations for a particular language (Chomsky, 1980b). At the initial state the child has the entire UG at his/her disposal, and is therefore capable of acquiring any language. In between the two states are a series of intermediate stages S_1, $S_2 \ldots S_n$. Acquiring a language means progressing from **So** to **Ss**. It is generally claimed that by the age of five or six, every normal child will have mastered most of the pronunciation of his/her native language (Salkie, 1990).

The rationale of UG lies in what Chomsky calls "the poverty of the stimulus". He claims that the evidence that the child receives from the environment is "degenerate" and "incomplete", and therefore insufficient to enable him/her to acquire certain rules of the mother tongue.1 Such evidence may be positive (e.g. actually occurring sound structures and sequences) or negative (e.g. explanations, corrections of wrong structures and sequences) (Cook, 1988). However, except for positive evidence, Chomsky claims that the child rarely encounters

1 This claim has however been shown to be unfounded. Evidence available from empirical studies in syntax (e.g. Sachs, 1977; Snow, 1976) indicate that the child is exposed to well formed language containing very few ungrammatical utterances.

negative evidence.2 He/she learns mainly from positive examples of what people say, and not from examples of what they do not say. Some structures may be so rare and marginal that the child obtains no exposure to them (Felix, 1984). In order to acquire those structures of the mother tongue for which there are neither positive nor negative evidence, the innate properties of language must be available to the child in the form of a UG.

The question remains as to what the relationship is between L1 and L2 acquisition in terms of the UG. Cook (1988) points out that L2 learners, unlike children acquiring their L1, do not begin from an absolute zero state (**So**). They already know a language, and therefore possess one instantiation of the UG principles. Nevertheless, not all learners have attained full competence (**Ss**) in their L1 by the time they begin to learn an L2. Child L2 learners (below the age of five), for example, are at the intermediate points between **So** and **Ss** in the L1, while adults have reached full competence. The varying degrees of knowledge of the L1

2 There may be indirect negative evidence, i.e. when the child notices the absence of a property in the input. The same may apply to L2 acquisition where the learner notices the absence of a property in the target language, which is present in the native language. However, in formal L2 learning situations negative evidence is often given in an explicit form, e.g grammatical explanations (which include rules of exception) and corrections. It is worth noting that studies on negative evidence in SLA have focused mainly on syntax (see for example, Schwartz and Gubala-Ryzak, 1993 and White, 1992).

possessed by the learners therefore constitute the initial state (**Si**) in L2 learning.

The final or steady state is not so clearly defined in the L2. In L1 acquisition this state is adult competence and is essentially complete, but the same cannot be said about L2 acquisition. Most people (except for a few balanced bilinguals) do not attain the same competence in the L2 as in the L1. As such, the steady state (**St**) that many L2 learners achieve differs from an L1 **Ss**, and varies from individual to individual (Cook, 1988).3

If the learner brings L1 instantiations to bear upon L2 acquisition, then transfer necessarily occurs. The main difference, therefore, between the mechanisms of L1 and L2 learning is interference. The learner is said to approach the L2 through the values of parameter settings for the L1. He/she may be forced to reset these parameters where they differ in the two languages. In some instances, the learner is unable to reset the parameters, and therefore approaches the L2 only through those for the L1 (cf.White, 1986).

The fact that the learner uses the L1 as a stepping stone to competence in the L2 implies that UG principles are available to him/her, indirectly, through the native language. However, the learner may have direct

---

3 It must be noted that the "steady state" of a bilingual does not equal the sum of two monolingual competences, but is something altogether different.

access to UG, without reference to the L1. Cook, 1988:185) states that,

> the learners' possession of an L1 is no guide in itself to languages that have other settings, unless UG is still present in their minds to prescribe the limits within which languages may vary.

In other words, the acquisition of new settings of UG parameters cannot be explained by the mere possession of an L1. In order to set new parameters for the L2, or reset L1 parameters in accordance with L2 constraints, the learner must have direct access to UG. The poverty of the stimulus applies to L2 in the sense that language acquisition is undetermined by the input data. Neither the L1, nor input from the L2 (i.e. L2 utterances) are in themselves sufficient to account for the acquisition of the language.

An alternative view holds that the learner may make no use of UG principles, directly or indirectly, in the learning process. This perspective therefore denies any parallels between L1 and L2 acquisition as far as access to UG is concerned.

In the discussion of the generative theory undertaken in this section we have considered the classical as well as UG versions, taking into account the added dimensions provided by non-linear phonology. However, there are other theories which represent a departure from the mainstream of generative phonology and

predate the non linear models. One of such theories is the natural phonology theory to which we now turn our attention.

3.4 Natural phonology theory

Natural phonology has its origins in David Stampe's (1969) dissertation (see also Donegan and Stampe, 1979). Though it represents a departure from the mainstream of generative phonology, it shares the basic assumption that innate universal principles underlie phonological development. However, it rejects the claim that human beings are endowed with a special language faculty. Rather, the theory postulates a universal innate system of phonological processes governed by forces implicit in human articulation and perception. These processes systematically map underlying phonological representations onto surface phonetic representations, and enable us to perceive in others' speech the phonological representations underlying such phonetic realizations (Donegan and Stampe, 1979:126). Phonological representation is defined in the tradition of Baudouin de Courtenay (1895) as the phonological "intention" of speech, i.e. what we intend to say.

Natural phonology posits that the entire system of universal processes is brought to bear upon language acquisition by the child, thus making him/her potentially able to acquire any language. This explains why in the early stages of first language acquisition, the processes

are manifest in their full blown form. However, as acquisition proceeds, the child gradually constrains those processes that are not applicable to the phonological system being learnt (i.e the adult system). For example, if acquiring a language in which nasalized vowels are distinctive, the child will have to constrain the relevant natural processes and thus achieve control of vowel nasalization (Clark and Yallop, 1990). A child learning L1 German will have to retain the process of terminal devoicing since German devoices all obstruents in final position. In a multilingual situation, the child sorts out the languages as well as processes, so that he/she possesses a different subset of the universal system for each language. Phonological development therefore consists in revising the universal system in the direction of a particular system or particular systems.

The mechanisms for this revision include, (a) suppression of a process - e.g the suppression of prevocalic tensing of vowels, flap deletion, (b) limitation, i.e the suppression of some part of a process by means of limiting either the set of segments or contexts to which the process applies, and (c) ordering of processes, resulting in the application of one process only (Donegan and Stampe, 1979). The mature system, therefore, consists of the residue of universal processes that are left after suppression, limitation and ordering.

Natural phonology claims that the child's mental representations correspond closely to the phonemic representations of the adult native speaker. This implies that the child's underlying representations remain constant as he/she gradually approximates the adult system. The process of revising the universal system towards the adult system of the L1 being learnt is often characterized by substitutions or what Stampe terms "phonetic changes". These arise from the child's failure to constrain certain processes at the early stages of acquisition, and accounts for the child's variable production. Consequently, the child deletes, devoices segments and inserts vowels in between segments. Given these facts, the child's substitutions are traced to productive rather than perceptual difficulties. Thus, as far as natural phonology theory is concerned, perception clearly precedes production (see section 2.6 above).

Phonological processes serve the communicative purpose of language. Donegan and Stampe (1979:142) identify three main types, each with distinct functions. The first type - prosodic processes - maps words, phrases and sentences onto prosodic structures (e.g syllabyzing, length and stress). The second type - fortition processes - serves to make segments more perceptible by intensifying salient features of individual segments and/or their contrasts with adjacent segments. They include diphthongization, dissimilations, syllabifications and epenthesis, and they apply mainly

to vowels in syllable peaks and to consonants in syllable onsets. They also apply in formal speech which require a high degree of attention to speech. The third type - lenition processes - makes segments more pronounceable by decreasing the articulatory distance between features of the segments and adjacent segments (e.g assimilations, monophthongizations, desyllabifications, reductions and deletions). They apply in syllable final positions, and in styles which do not require a high degree of attention to speech, i.e. casual speech.

However, not all phonological alternations are governed by natural processes. Alternations involving, for example, velar softening and the umlaut rules of Germanic languages, are governed by morphonological and morphological rules (Kolaczyk, 1987). Natural phonology therefore distinguishes processes, which are phonetically motivated, from rules, which lack such motivation. Rules have to be learnt as a matter of convention. Donegan and Stampe (1979:144) claim that we pronounce "profound" as [profãʊ̃nd] and not as [profoʊ̃nd] because we cannot say the latter without acquiring greater velar precision. However, "profundity" is pronounced as [profʌ̃nditi] instead of [profãʊ̃nditi] by convention - because that is what other speakers say.

Processes are said to apply involuntarily and unconsciously, while rules apply voluntarily and consciously, being formed through the observations of linguistic differences among speakers (p. 144). Rules

unlike processes are open to exceptions or morphological conditioning. Processes represent constraints on pronunciation and can only be violated if the learner makes a special effort. Processes apply to slips of tongue and are optional, while rules are obligatory, and are sometimes the result of fossilization. As for ordering, rules are said to apply before processes, fortitions before lenitions. Finally, it is residual processes and not rules which are extended to the pronunciation and perception of foreign words.

Donegan and Stampe (1979) claim that their theory is "natural" in that it presents language not as a conventional institution, but as a "natural reflection of the needs, capacities and world of it users" (1979:127). The theory seeks to present language as following from the nature of things, and is therefore not intended to generate a set of phonologically possible languages in the Chomskyan sense. Donegan and Stampe reject structural and generative phonologies on the grounds that they are not "true" theories given that neither is genuinely open to falsification by data (Clark and Yallop, 1990). Natural phonology, nevertheless, fails on its part because it makes no allowance for hypothesis testing and creativity on the learner's part. As such a particular phonological system, or the acquisition of such a system, is the way it is, because of purely linguistic constraints imposed by the nature of the human language

and human articulatory and perceptual systems (Macken and Ferguson, 1987).

In spite of this rather deterministic approach, and in spite of the fact that Stampe claims very little about foreign accent, natural phonology theory has implications for L2 phonological acquisition. The theory has been explored in L2 acquisition by Major (1987b) and Kolaczyk, 1987, 1990). According to Major, the similarity between L1 and L2 phonological acquisition lies in the fact that both involve the elimination or constraining of processes which are not characteristic of native adult pronunciation. In L1 acquisition this involves constraining a set of universal (developmental) processes. In the L2 it involves constraining the same set of universal processes, but only some of which have been retained (due to already completed L1 acquisition), and interference processes which may or may not be the same as these developmental processes. In other words, processes which are not characteristic of the L2 system are eliminated.

Given the mechanism of elimination, the L1 knowledge that the learner brings to bear upon L2 acquisition consists of a much reduced system comprising selected processes and underlying representations together with learned rules (Kolaczyk, 1987, 1990). The L2 learner's phonological "intentions" at the early stages of acquisition resemble mainly those of the L1. As acquisition proceeds, the learner gradually eliminates

unwanted developmental and interference processes in order to achieve adult L2 pronunciation. The elimination process may not always be successful if fossilization sets in in the course of acquisition.

Natural phonology predicts that unsuppressed L1 processes will appear in L2 acquisition. However, not all such processes will be observed. We noted in chapter 2 that consonant harmony is strictly an L1 process which does not occur in L2 acquisition. Furthermore, not all L1 processes will surface with every L2 learner. Fast learners may constrain non-native processes so that they do not surface. L1 processes which appear may or may not resemble those required for L2 acquisition. The subconscious application of native processes to L2 strings will result in interference unless a native process is identical with the one selected to operate in the L2 (Kolaczyk, 1987). For example, terminal devoicing is not an interference process in the IL German of Polish speakers, and vice versa, because it occurs in both languages. However, the processes of vowel nasalization and gliding which occur in Polish but not in English cause interference in the IL English of Polish speakers (Kolaczyk, 1987). So, Polish learners will have to eliminate these processes from their IL English. If on the contrary, the native language system lacks some process operating in the L2, such a process is learnt in the same way as L2 rules.

With reference to the last point, it can be argued that the non-L1 process which operates in the L2 is not entirely new after all. It is possible that certain processes are not completely suppressed during L1 acquisition, but lie latent in the learner's brain. Therefore when learning an L2 that requires a non L1 process (supposedly eliminated during L1 acquisition), the learner simply "reactivates" such a process from latent memory. The question is, which processes are reactivated? Those which originate from process types which are not completely suppressed in both languages? (Kolaczyk, 1987). If we accept that certain universal processes are "latent" and can be reactivated when the need arises, then we can argue that the set of processes still available to the learner is larger than claimed. It comprises not only the residue that have survived suppression, but includes other processes which have not been completely suppressed. We can refer to the latter as latent processes.

One major difficulty, however, with natural phonology is the lack of distinction between interference and developmental processes in L2 acquisition. Stampe seems to view both processes as one and the same, thereby denying the well established notion that certain substitutions in the L2 are attributable to the L1 of the speaker (e.g the substitution of [x] for English [ɹ] by Brazilian Portuguese speakers), while others are attributable to developmental processes (e.g the use of

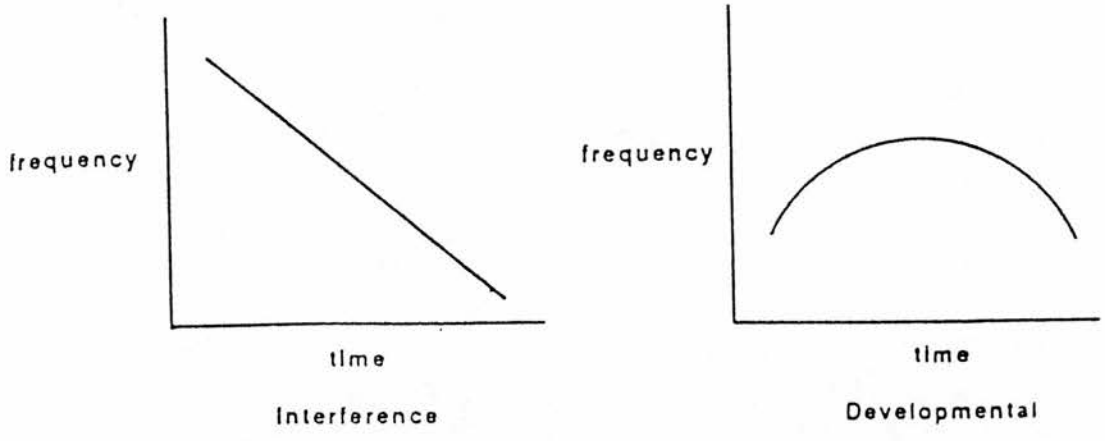
the Schwa paragoge by Brazilian Portuguese speakers) (Major, 1987b). The relationship between the two processes are explored in Major's Ontogeny model which we examine below.

3.5 The Ontogeny model

The Ontogeny model proposed by Major (1986b, 1987a, 1987b) describes the diachronic as well as synchronic relationship between transfer and developmental processes. Chronologically, the model claims that at the initial stages of L2 acquisition interference predominates at the expense of competing developmental processes, thus preventing the latter from surfacing. During the course of acquisition interference decreases while developmental processes increase and decrease over time. Finally, as the learner's pronunciation becomes more target-like, both processes are eliminated. In the acquisition of English **dog** by Brazilian Portuguese speakers, for example, the following pattern is generally observed: [dɔgi] (interference) > [dɔgə] (developmental) > [dɔk] (developmental) > [dɔg] (Major, 1987a). The [i] paragoge produced after the voiced stop /g/ is an interference process from Brazilian Portuguese. Later [i] is replaced by a Schwa paragoge [ə], which is a developmental process. Another developmental process - terminal devoicing - follows, before the learner eventually acquires the target pronunciation. The pattern shows the gradual elimination

of transfer and developmental substitutions as the learners acquire target pronunciation. The emphasis here is on the elimination of processes in accordance with the principles of natural phonology. Graphically, the relationship between processes is represented as follows:

Figure 3.1: Relationship of interference and developmental processes to time



Major (1987a:103)

The reason why interference processes predominate at the early stages of acquisition is because the learner has mastered very little of the L2, and therefore relies much more on his L1. His/her IL system shows a high frequency of interference substitutions. As learning progresses, the IL system becomes increasingly affected by the L2 causing new structures (i.e developmental substitutions) to appear, while old L1 structures begin

to decrease in frequency. In other words, as the IL approximates the target, the learner's substitutions follow naturally from the L2 rather than the pure L1 system. The suppression of interference processes allows developmental processes to operate. Eventually both transfer and developmental processes are eliminated. Nevertheless the relationship between interference and language proficiency is not as direct as it seems. Sometimes interference does not occur until after the learner has reached a certain level of proficiency in the L2 (Odlin, 1989). Interference may persist in the form of fossilized structures even with increased L2 knowledge.

Interference processes may be replaced by a series of developmental processes, consecutively, until the target pronunciation is reached. For example, in the Brazilian Portuguese speakers' acquisition of English **dog** referred to above, the interference process ("i" paragoge) is replaced by two developmental processes (Schwa paragoge and Terminal devoicing) respectively, before the target is achieved. As already indicated above, it is not always the case that every learner will achieve native pronunciation, as fossilization may set in.

The Ontogeny model is intended to be an abstraction of the general pattern of L2 learners, and does not claim that acquisition will always proceed as described. In other words, acquisition will not necessarily be gradual. It may follow discrete jumps so

that the learner bypasses certain stages for a given phenomenon and reaches the target faster - e.g a learner who correctly produces English [ɹ] on the first attempt. This will depend on other factors such as differences in learning abilities, inherent variability of phenomena in general, and the L1 of the learner. A good learner will achieve native pronunciation faster than a poor learner. Different tasks will yield different accuracy rates for target production. Speakers of some L1's may bypass certain processes in the acquisition an L2 phenomenon. For example, a speaker of Chinese, a language with no final obstruents, will either devoice them in English utterances or use the schwa epenthesis. No clear L1 interference is observed and the two processes cited are developmental since neither occurs in Chinese.

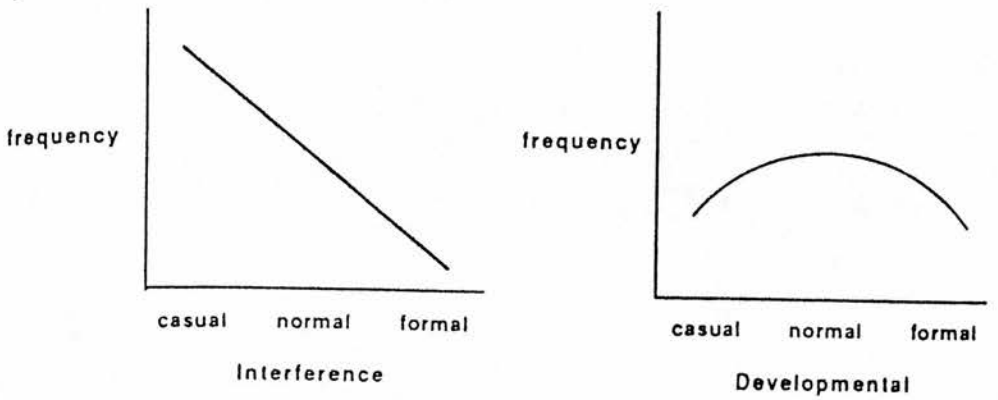
While the model claims a hierarchical relationship between transfer and developmental processes, it does not specify the relative frequency of these factors (Major, 1986b). As such, we do not know the exact percentage of increase or decrease of both factors in relation to target production. Major suggests that target production will remain constant if transfer decreases at the same rate as developmental factors increase. On the contrary, if transfer decreases at a faster rate than developmental factors increase, then correct production will necessarily increase. This would imply that more transfer than developmental substitutions are replaced by

target segments. However, in an unlikely situation where transfer decreases at a slower rate than developmental errors, target production is predicted to decrease. We can, therefore, conclude from the above statements that an increase in target production depends more on the rate of decrease of transfer processes than on the rate of increase of developmental processes.

The relationship between transfer and developmental factors also holds for stylistic shifts in that as style becomes increasingly formal, e.g conversation --> text --> sentence --> word, transfer decreases and developmental factors increase and then decrease. The relationship is represented in Figure 3.2 below. Target production is more likely to occur in formal word list reading than in free speech. L1 interference is predicted to be stronger in the latter than in the former because the phonostylistic processes of casual speech are less constrained than those of formal speech (Kolaczyk, 1990). The casual style demands less attention to speech than the formal style. In "normal" styles (i.e. styles in between the two extremes of the continuum) interference decreases while developmental factors increase, and then decrease in the more formal styles. This suggests that the speaker is able to suppress processes in certain speech styles, which apply inadvertently in other styles.

Developmental processes are governed by universal stylistic considerations involving fortition and lenition as described by Donegan and Stampe (1979). Thus,

Figure 3.2: Relationship of interference and developmental processes to style



Major (1987a: 107)

epenthesis (a fortition process) is predicted to occur more frequently in word list reading, and consonant deletion (a lenition process) more frequently in conversation. The high level of attention required in formal situations triggers fortition processes, while the low level of attention in casual situations triggers lenition processes.

Major says nothing about the exact frequency of transfer and developmental processes relative to target production along the style continuum. We may then ask if the rate of decrease of transfer and the rate of increase of developmental processes affect the frequency of target production. It is not quite clear how this can be determined synchronically along a style continuum. One issue which the model does not address is that of markedness.4 It is to this phenomenon that we **return** our attention.

3.6 Phonological markedness

The term "markedness" is by now a rather familiar one: the reader will recall the numerous references that have already been made to it in the preceding and present chapters. However, given the importance of the concept in phonology, it is only logical that markedness is discussed as a theory in its own right.

Markedness describes an asymmetrical relationship between two elements, where one is regarded as basic, expected, natural and therefore unmarked, and the other as unexpected, complex, unnatural and therefore marked (Cairns, 1986; Gair, 1988; Houlihan and Iverson, 1980; White, 1986). This concept has been of central importance to most approaches to phonological theory, and the

4 However, see Major (1987c).

criteria for defining markedness have differed accordingly.

The original Prague School notion of markedness owes its existence to the phenomenon of neutralization. According to Trubetzkoy (1939), segments in opposition are considered to be "marked" and "unmarked" if the opposition is neutralizable in certain contexts. Neutralization affects privative as opposed to non-privative oppositions. The former are characterized by the presence or absence of a phonetic feature, given a + or - value (e.g voicing, aspiration and nasality). (Cairns, 1986). In general, it is the unmarked member of an opposition that occurs in the position of neutralization. Voice contrast is neutralized word finally in German, such that only voiceless sounds - [p], [t], [k], [ʃ] and [s] - occur in that position (Hyman, 1975). Hence, voicelessness is unmarked and voicedness marked word finally.

The problem with Trubetzkoy's concept of markedness is that it is restricted to individual languages. As such, markedness relations between segments differ from language to language, and are not generalizable. An opposition may neutralize to one segment in language A, but to the other in language B. The resultant indeterminacy makes the distinction between marked and unmarked segments difficult. This problem is sorted out when markedness is treated as a universal rather than as a language-specific notion.

The concept of universal markedness was introduced by Jakobson (1968), and latter adopted and modified by Chomsky and Halle (1968). Basically, it derives from considerations regarding the relative frequency of phenomena across languages, complexity, naturalness, order of acquisition, typological universals, language loss and change (Chomsky and Halle, 1968; Herbert, 1986; Greenberg, 1965, 1966; Jakobson, 1968). Some of these notions were discussed within Jakobson's framework in section 3.2., and are only reconsidered here for purposes of clarification and extension to other frameworks.

Markedness within the generative framework has been partly examined in terms of parameter settings for UG (see section 3.3). However, we are concerned here with the formal model of markedness presented in the Standard (SPE) Theory. In the said model markedness establishes a relationship between relative frequency, naturalness and complexity. The more frequent a segment type, process or system across languages, the more natural and less complex it is considered to be. Low markedness values are assigned to such segment types, systems or processes. On the contrary, those which are less frequent, less natural and more complex are assigned high markedness values.

This view of markedness follows from a reconsideration of the phenomenon of naturalness which was initially defined by a feature-counting metric. Natural classes of sounds were those characterized by relatively few features, and less natural classes by an

increased number of feature specifications. Hence, obstruent consonants described as [-sonorant] were regarded as more natural than voiced consonants (other than laterals) specified as [+voiced, +consonantal, -lateral] (Clark and Yallop, 1990). However, the evaluation of naturalness in this manner was riddled with flaws since the occurrence of a limited number of features for particular segments did not necessarily entail a more natural class of sounds.

Chomsky and Halle were aware of this problem, and proposed a remedy which took into account the "intrinsic values" of various feature specifications. Certain combinations of features were expected (or natural), while others were rare or impossible (Kaye, 1989). They proposed that the binary values, + and -, be replaced by the symbols "m" (marked) and "u" (unmarked) respectively. The new values are considered to reflect expectedness or naturalness, and are converted to + and - values by marking conventions. E.g the unmarked value for the feature "back" [*u*back] is [+back] if the vowel is [+low] (Hyman, 1975:148). The universally unmarked low vowel is /a/. The unmarked value for the feature "round" [*u*round] is [+round] if the vowels are both back and non low (e.g. /u/). The unmarked value for voicedness [*u*voiced] is [-voice] if the consonants are obstruents. Unmarked values are cost free, while all others cost 1. The markedness of a segment is, therefore, derived by

adding up the cost of features. Less marked segments cost less, while more marked segments cost more.

In addition to marking conventions, Chomsky and Halle propose the notion of linking which allows marking conventions to monitor phonological rules. As such naturalness conditions are not restricted to segments but are also extended to rules. For example, the rule which converts /i/ to [u] as in:

$$\text{(a)}\ \underset{V}{[+\text{high}]} \longrightarrow \begin{bmatrix} +\text{back} \\ u\text{round} \end{bmatrix}$$

is considered to be more natural than the one which changes /i/ to [ɯ] as in:

$$\text{(b)}\ \underset{V}{[+\text{high}]} \longrightarrow \begin{bmatrix} +\text{back} \\ m\text{round} \end{bmatrix}$$

This is because the unmarked value for "round" for non low vowels (i.e. [+round]) is identical with the specification "back" (Hyman, 1975;154/155). Rule (a) links up with this marking convention and converts /i/ to [u]. In the case of rule (b) the marked value for "round" for non low vowels (i.e. [-round]) is not identical with the specification "back". Linking, therefore, specifies that "when a feature changes in a rule, all other features whose markedness values depend on the changed feature take on their unmarked values with no cost" (Lass, 1984: 197). Note that rule (a) costs 3 features as opposed to rule (b) which costs 4 features.

Recall that unmarked features are cost free. Hence [*u*round is not counted for rule (a). On the contrary, marked features cost 1. Hence [*m*round] is counted for rule (b).

This view of markedness has, however, been rejected on the grounds that the formalism embodied in the evaluation of naturalness/markedness fails to do justice to naturalness itself. The arguments put forward for the justification of this evaluation are therefore said to be vacuous, circular and non explanatory (Allan and Bauer, 1991; Clark and Yallop, 1990; Lass, 1984;).

With regard to language acquisition, evidence shows that first language learners generally acquire unmarked sounds earlier and more easily than their marked counterparts (Jakobson, 1968; Menn, 1986). Therefore, acquisition follows the invariant order "unmarked" before "marked". Children also tend to replace marked sounds with unmarked ones during the early stages of acquisition. In aphasia unmarked sounds are lost before marked ones, and in sound change unmarked sounds replace marked ones (Jakobson, 1968; Hawkins, 1984).

The question as to whether L2 acquisition strictly follows the universal order was partly answered in section 3.2. We observed that the L1 plays a major role in determining the sequence of acquisition of phonological elements. The degree of influence of the L1 may be further determined by the markedness differential between the native and target languages. Thus, the claim

that, developmentally, L2 acquisition proceeds from unmarked to marked (Hyltenstam, 1984), is true only under certain conditions. For example, when the L1 has the unmarked form and the L2 the marked form, then the unmarked form may appear in early IL and the marked in later IL (cf. Eckman, 1977). However, if the L1 is more marked than the L2, two views are proposed: either learners follow the universal order, or they use the marked form at early and late stages of acquisition. White (1986) opts for the latter view. She postulates that where the learner is unable to unset a marked L1 parameter setting, he/she will transfer the marked setting into the L2 even if the latter has an unmarked setting.

Predictions about the difficulty and directionality of acquisition often stem from typologically-based universal implications of the Greenbergian and Jakobsonian types. These universals are derived from a survey of surface patterns in a wide variety of languages and the implicational relations which hold between them. The markedness relations derived from universal implications are such that,

> A phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B; but the presence of B does not imply the presence of A (Eckman, 1977:320).

The following are some examples of implicational universals:

(a) The presence of voiced stops implies the presence of voiceless stops but not vice versa (Jakobson, 1968:70).

(b) The presence of voiced fricatives implies the presence of voiceless fricatives but not vice versa (Jakobson, 1968:70).

(c) The presence of fricatives implies the presence of stops, but not vice versa (Jakobson, 1968:51).

(d) The presence of syllables containing sequences of *n* consonants implies the presence of syllables containing the sequences of *n-1* consonants (in corresponding positions) but not vice versa (Greenberg, 1965:5).

The following markedness relations are derived from the above universals:

(a) Voiced stops are marked with respect to voiceless stops.

(b) Voiced fricatives are marked with respect to voiceless fricatives.

(c) Fricatives are marked with respect to stops.

(d) Consonant clusters are marked with respect to singleton consonants.

The interaction of universal implications with language acquisition constraints form the locus of Eckman's Markedness Differential Hypothesis to which we turn our attention.

3.6.1 The Markedness Differential Hypothesis

Given the above definition of markedness, Eckman (1977) proposes a Markedness Differential Hypothesis (MDH) which states that,

> The areas of difficulty that a language will have can be predicted on the basis of a systematic comparison of the grammars of the native language, the target language and the markedness relations stated in universal grammar, such that,
>
> (a) Those areas of the target language which differ from the native language and are more marked than the native language will be difficult.
>
> (b) The relative degree of difficulty of the areas of the target language which are more marked than the native language correspond to the relative degree of of markedness.
>
> (c) Those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult.
>
> (1977:321).

The MDH claims that given a number of crosslingual differences, and therefore given also the number of potential areas of difficulty for learners, only some differences will result in difficulty, while others will not. Moreover, areas of difficulty will be predictable on the basis of a markedness differential between L1 and L2. Where the L2 exhibits a more marked form than the L1 difficulty will arise. Where the L2 is less marked than the L1 (including instances where no markedness relations

exist between the L1 and L2), the MDH predicts no problems.

The MDH is an improvement upon Lado's (1957) Contrastive Analysis Hypothesis (CAH) in that it predicts difficulty based on cross-lingual differences by incorporating universal grammar principles. Thus, as already noted above, only certain types of difficulty will occur, and only such difficulties will cause negative transfer. That is, where the L2 is more marked than the L1, transfer will occur. Consequently, unmarked categories from the native language will be substituted for corresponding marked categories in the target language (cf. Cook, 1985; Hyltenstam, 1984; Kellerman, 1983; Rutherford, 1982). No transfer effects are predicted where the L2 is less marked than the L1.

Eckman illustrates this using the acquisition of voice contrast by German learners of English, and English learners of German. Both languages have voice contrast initially and medially. The difference between them lies in the maintenance/ non-maintenance of a voice contrast in obstruents finally. While English has both voiced and voiceless obstruents in this position, German retains only the voiceless counterpart. As already indicated above voiced obstruents are more marked than voiceless obstruents. In addition the maintenance of a voice contrast is more marked word finally than medially, and more marked medially than initially. The maintenance of a contrast medially does not imply such a contrast finally,

and the presence of a contrast initially does not imply a contrast medially and finally.

Given that the position of contrast and non contrast in English and German respectively is the most marked, and that German exhibits only the unmarked voiceless obstruents in this position, the MDH predicts that German speakers will have difficulty acquiring the more marked voiced obstruents finally in English. They will consequently substitute unmarked voiceless forms in the said position. However, English speakers will have no difficulty suppressing the most marked word final L1 distinction when learning German. While German speakers need to master a more marked system, English speakers need to master a less marked system.

However, the MDH has been criticized on several grounds. Its ability to predict and explain all SLA phenomena including those which exhibit no subordinate markedness relations has been questioned (Altenberg and Vago, 1987; Hammarberg, 1988). Eckman (1985) has rejected this criticism, arguing that the MDH makes predictions about all aspects of the L2 and L1, as well as those aspects which have no markedness relations. The MDH is also criticized for capturing too few markedness relations by relying on a single criterion - implication (Hammarberg, 1988). Kean (1984) considers it implausible that learning implicationally more marked elements should pose more difficulty than learning implicationally less marked elements (cf. White, 1986).

The outcome of these criticisms is the call for a more viable model of markedness. Fellbaum (1986) advocates the incorporation of the concept of contextual/ distributional relations of segments in the MDH (see Fellbaum, 1986). Hammarberg (1988) suggests that the notion of crosslingual markedness differential be dispensed with and that the degree of difficulty in L2 acquisition be correlated with the relative degree of markedness in the L2 itself.

A "middle-of-the-road" proposal by Benson (1986), however, sounds more plausible than Hammarberg's. That is, that cross-lingual markedness relations must be considered together with markedness relations within the target language in order to predict and explain learner-difficulties. First of all, an overall markedness relation is established between the L1 and L2 for a given phenomenon. For example, an L2 which allows consonant clusters is more marked than an L1 which allows only singleton consonants (Greenberg, 1965). Speakers of the less marked language will have difficulty acquiring the more marked language with respect to the given phenomenon. However, the degree of difficulty may depend on the sub-levels of markedness which exist among members of the given phenomenon in the target language: not all clusters are equally marked.5 Relative degrees of markedness among clusters in the L2 must therefore be

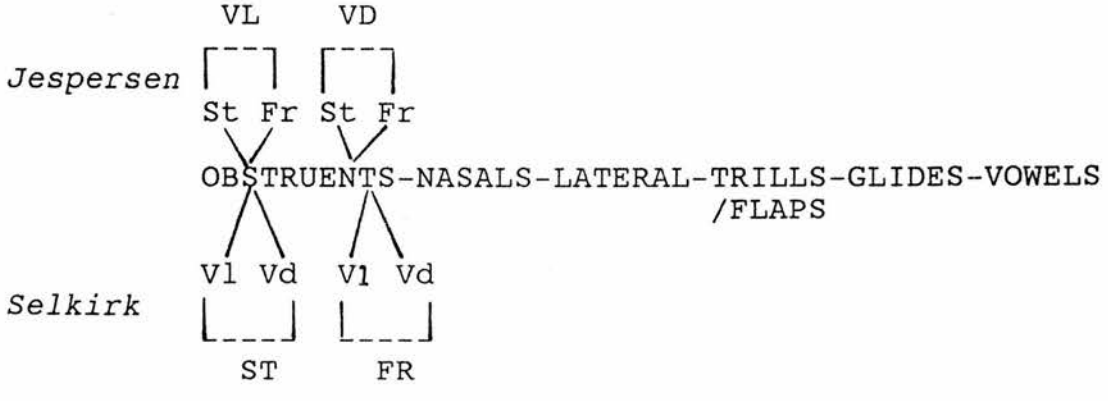
5 See the discussion of Benson's (1986) study of Vietnamese learners of English in Chapter Two.

established. One way of achieving this is by considering the theory of sonority relations.

3.6.2 The Sonority hierarchy

In section 3.3 we observed that one of the few principles proposed specifically for phonology is the Sonority Principle. This principle describes the universal characteristic of the intrinsic syllable structure which is that segments are ordered hierarchically from the peak (i.e. the vowel) to the peripheries in decreasing order of sonority (Hooper, 1976; Jespersen, 1904; Saussure, 1916/59; Selkirk, 1984). This is also referred to as the Sonority Sequencing Generalization (SSG - Broselow and Finer, 1991), and may be represented as follows:

Figure 3.3: The Sonority Hierarchy



(NB: VL = voiceless; VD = voiced; ST = stop; FR = Fricative)

According to the hierarchy, the vowel constitutes the most sonorous segment in the syllable. Consonants are less sonorous relative to vowels, with obstruents being the least sonorous and glides the most sonorous. However, there seem to be different views regarding the sonority of obstruent categories. In the earlier version of the hierarchy proposed by Jespersen, the feature voice is given prominence over others like stridency (i.e. acoustic intensity), strength (weak/strong) and stricture (complete/partial obstruction of airflow which serve to distinguish obstruent plosives from fricatives).[6] The reverse is the case in the version proposed by Selkirk. These differences are reflected on the left side of the hierarchy.

The Sonority Principle restricts the order of consonants in syllable onsets, while the number and type of consonants in onsets are determined by the distance in sonority between adjacent segments (Broselow and Finer, 1991). Languages may vary in the degree of differences in sonority required between adjacent segments (i.e. in the parameter settings for sonority distance). In the case of clusters, the general pattern is that consonants closer in sonority are less likely to form cluster sequences

6 Given these features, less sonorous obstruents require more acoustic energy, are stronger and involve complete obstruction of airflow. Stops fall into this category. More sonorous obstruents require less acoustic energy, are relatively weak and involve partial obstruction of airflow. Fricatives fall into this category.

than those which are farther apart in sonority. This is known as the Minimal Sonority Distance Parameter (MSDP). In English, the plosive may combine with liquids and glides, but not with nasals. Plosives and liquids/glides are farther apart in sonority than plosives and nasals. An exception to the rule is the combination of the fricative "**s**" with nasals (segments close in sonority), as well as with liquids (segments wide apart in sonority).

Implicational relations may be derived from the hierarchy such that the presence of clusters whose members are closer in sonority (A) implies the presence of those whose members are farther apart in sonority (B). Therefore, a language with clusters of plosive + liquid sequences also has clusters of plosive + glide sequences. According to the laws of markedness, type A clusters are more marked than type B clusters. It is predicted that the more marked clusters will be harder to learn than the less marked clusters. The speaker of an L1 which allows no clusters (e.g. Ibibio), acquiring an L2 which does (e.g. English) will have difficulty only with certain types of clusters in the L2, according to the markedness principles stated above.

In this section we have considered the phenomenon of markedness from a general perspective. Particular attention has been paid to the MDH proposed by Eckman. It is argued that the markedness relations as conceived in the MDH do not account for all L2 phenomena. One way of

resolving this problem is by considering the sonority relations between phonological elements. Another way of explaining L2 phenomena is to consider the sociolinguistic dimensions of language. This forms the focus of the next section.

3.7 The Labovian Model

Of the models which deal with the sociolinguistic aspect of language acquisition, the most influential appears to be the Labovian model. Based on William Labov's (1969, 1970, 1972) studies of native speaker speech, the model is intended to account for systematic variability in language acquisition with reference to factors such as speech situation, discourse topics, tasks. Its central claim is that variability results from style shifting which in turn results from fluctuations in the degree of attention paid to form. This claim is embodied in the axioms of Labov's Observer's Paradox which state that,

(a) there are no single style speakers: people adapt their speech to suit the social context;

(b) styles are ranged along a continuum, measured by the amount of attention paid to speech;

(c) the vernacular style demands the least attention to speech; it exhibits the most systematic and internal consistency, while other styles show variability;

(d) it is not possible to tap the vernacular

style through the systematic observation of the speaker;

(e) the only way to obtain good data on the speech o the language user is through systematic observation.

(Labov, 1972:208, 209)

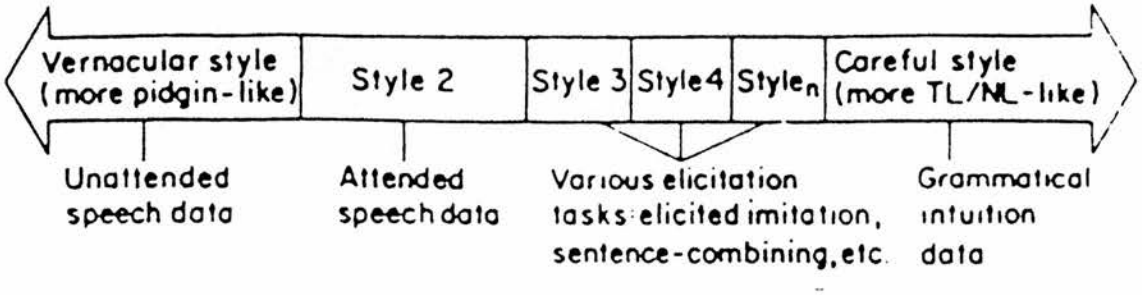
Variable behaviour is claimed to be the outcome of an underlying heterogeneous competence, which can be accessed through a variety of tasks differing in levels of formality.

In SLA, the Labovian model is exemplified in the works of the Dickersons (1975, 1976, 1977), Tarone (1979, 1982, 1983, 1988) and Beebe (1987a) (see section 2.5 above). The model is represented in the form of a Capability Continuum which Tarone (1983, 1988) proposes to account for variability in the IL system. She argues for the superiority of the Continuum paradigm over the Chomskyan paradigm (cf. Adjemian, 1976) or Krashen's Monitor Theory. Tarone uses the term "capability" as opposed to "competence" because the latter refers to "the sort of linguistic knowledge...accessible almost in its entirety to a form of introspection in that grammatical judgements may provide the linguist the most accurate data on it" (1983:151). That is, "competence" describes purely linguistic (grammatical) knowledge, the kind Chomsky (1965, 1980) claims is the prerequisite of the ideal speaker-hearer. Within the Chomskyan paradigm competence is assumed to be homogeneous, and is

distinguished from performance (behaviour) at which level variability is said to occur.

"Capability", on the contrary, refers more broadly to "that which underlies all regular behaviour", that is "regularities in learner production and perception... as well as in making judgements on grammaticality" (Tarone, 1983:151). The kind of "competence" which Tarone envisages therefore, is a heterogeneous one, and includes grammatical knowledge. The heterogeneous IL Capability is made up of a continuum of overlapping styles which include careful, vernacular as well as intermediate styles. These styles are related in terms of variable and categorical rules. The Capability continuum is represented below.

Figure 3.4: Capability Continuum



Tarone (1983:152)

A learner's performance at any time may be located anywhere along the continuum depending on how much attention is paid to form at that instance. A superordinate (careful) style is produced when the speaker pays more attention to form. This style exhibits more target language as well native language forms, and is thus highly variable. The vernacular style occurs when the speaker pays minimum attention to form. This is the style least invaded by the target or native languages, and is therefore more systematic and internally consistent. It is unique to the IL in that it contains structures which are developmental or IL approximations, i.e. structures which "share features in common with pidgins, early child languages, and other 'simplified' languages" (Tarone, 1983: 155).

With regard to the place of the native language in the continuum of styles ,Tarone's views differ slightly from those of Dickerson (1974, cited in Beebe, 1988). While Tarone claims that L1 forms will occur in the formal style together with target forms, Dickerson claims that such native language forms will be found in the vernacular style. However, it seems to be the case that those L1 variants which persist in the formal style are socially prestigious variants which are transferred to equivalent settings in the L2 (cf. Beebe, 1987a; Schmidt, 1987). When learners attend to speech in formal styles, they not only produce a high incidence of L2 forms, but

also a high incidence of L1 forms if these are associated with formal use in the native language.

The learner's speech varies systematically with elicitation tasks. The grammaticality judgement task, for example, provides information about only the careful style. Data from this style also include those obtained in structured interviews with unfamiliar researchers. The reading of word lists or connected sentences tap styles occurring intermediately along the continuum, while the vernacular style is elicited in conversations. Tarone stresses the need to collect data from as many styles as possible to allow for the accurate identification of regularities within and among them, as well as the influence of the L1, L2 and pidgin-like structures (1983: 155).

Tarone (1988) envisages two ways in which new forms may be incorporated into the IL: (a) forms may be spontaneously produced first in the vernacular style, or (b) they may appear first in the most formal style, and gradually spread over time into less formal styles. She claims that new forms are more likely to spread as described in (b) than in (a). At the same time, she acknowledges that forms may also spread from the vernacular style into more formal styles. Thus, there is some remote possibility for a bidirectional spread.

The means of spread is presented in the works of the Dickersons (1975, 1976, 1977) which we have considered in detail in section 2.5. The process is

described in terms of a wave mechanism (or ordered decomposition) which is equated with sound change in the speech of native speakers of a language (Dickerson, 1976). According to the wave mechanism, change occurs in small increments along three linguistic dimensions. Firstly, the change affects the set (or subset) of words containing a particular word class (e.g. /l/ word class). Secondly, the affected subset of words takes a phonetic shape or variant which is slightly different from the original variant but in the direction of the target variant. Thirdly, each subset of words in which the new variant appears shares a particular phonetic environment. The central claim of the model, therefore, is that acquisition of sound segments proceeds by the movements of variations within environments of variable word classes.

The learner initiates change in one linguistic environment which gradually spreads to other environments in an ordered sequence. This order is maintained systematically over time. The acquisition process may begin with the appearance of the native language variant in all environments, followed by that of the target variant in one environment (cf. Gatbonton's 1978 Diffusion Model), or by the appearance of variants in any or all linguistic environments. What is worth noting is that the order of environments in terms of favourability to the new variant remains constant across tasks and styles. Tarone claims that evidence from the

Dickerson studies support her proposed order of acquisition, which is that, developmentally, target forms get to the vernacular via the careful style. Those IL structures which appear spontaneously in the vernacular style are gradually replaced by the target variants seeping in from the more formal styles. At any point in time the careful style exhibits the highest proportion of target variants, with less and less formal styles eliciting progressively fewer target variants.

The Labovian model has been criticized on several grounds. The lack of a distinction between underlying linguistic competence and IL behaviour (or performance) has been questioned by Gregg (1989), who working within the Chomskyan paradigm insists that stylistic variability is not an integral part of competence, but of performance. The latter is not "acquired" in the sense that the former is. Variability occurs at the performance level as a result of the invasion of the learner's underlying homogeneous system by other rule systems (e.g. the L1 or IL rules) (Adjemian, 1976). An adequate theory of SLA need not be concerned with performance (i.e. variability), but with the acquisition of linguistic competence. Tarone (1988) argues on the contrary that if other linguistic systems are able to "invade" in the said manner, then they must be known to the learner, and must be part of his/her linguistic knowledge. Therefore any observable variable behaviour must be an integral part of

linguistic competence, and should be accounted for by any theory of second language acquisition (Tarone, 1984).

Perhaps the main criticism of the model has been directed at its explanatory power, or the lack of it. Bialystok and Sharwood-Smith (1985), for example, draw attention to what they consider to be the model's lack of specificity regarding the psychological processes that underlie IL variability. The fact is that Labov does cite a cause for style shifting, in terms of the psychological process of attention to speech (Tarone, 1988).

However, the question remains as to whether the construct of "attention to speech" constitutes an adequate explanatory factor for IL variability. Beebe (1982) and Bell (1984) contend that attention to speech is not the sole cause of style-shifting, or even the single style dimension underlying all style shifts. Learners may equally shift styles as a result of socio-psychological variables, e.g. ethnic identity, solidarity, topic expertise, the status of the participants (Beebe, 1982) and task related factors such as task complexity or time constraints (Ellis, 1985a). In other words, style-shifting may result from an array of factors, which may or may not include attention to speech. At the same time, the possibility that a single factor may indeed contribute to style-shifting, or at least play a dominant part in the process, is not completely ruled out.

The difficulty is in deciphering which of the factors is (or are) responsible for style-shifting. This problem is further compounded by the fact that certain variables, e.g. attention to speech, do not lend themselves to empirical measurement. It is difficult to measure the degree of attention paid to speech in different tasks. We cannot be sure that the learner is paying more attention to form in word list reading than in sentence reading or story-telling. At the same time we cannot be sure that the same degree of attention is maintained within a given task. Factors such as the nature of the task items or fatigue must be considered. For instance, in a casual style, attention may increase for an informationally loaded word, or in a formal style, it may decrease for a frequently used word (Kolaczyk, 1990). A task that is particularly long will not always elicit the same degree of attention throughout its duration. The setting in which L2 acquisition takes place is another factor to consider. Ochs (1979) and Rampton (1987) (both cited in Makoni (1989) posit that learners who have acquired the L2 in a predominantly formal environment require a lot of attention to speech in order to construct a vernacular style.

Attention to speech may not always correlate positively with standardness or correctness, in that the formal style will not always elicit target-like production. Tarone (1982, 1988) herself, has clearly demonstrated that the formal style can display native

language variants as well, and therefore lower correctness. We have already considered one explanation for the persistence of native language variants in the formal style: such variants may be socially prestigious in the L1. Alternatively, they may occur as fossilized structures in the learner's IL system such that no amount of attention paid to speech will prevent them from surfacing in the formal style.

In as much as we do not dispute the claim by Tarone (1983) and Dickerson (1974) that less target-like production occurs in the casual than in the careful style, we would like to agree with Sato (1985) that, in certain instances, the vernacular style does evidence more correctness than the formal style. The reason is that while formal tasks do require a great deal of attention, this attention is not simply paid to language form, but also to "other demands on real time discourse production: recording and encoding of rhetorical structure, lexical items, clause sequencing" (Sato, 1985:195).

The "attention to form" paradigm is problematic in another sense. The learner may pay attention to form in order to realize specific social goals (Makoni, 1989). A learner seeking to project the status of a non native speaker will tailor his/her speech away from that of the interlocutor, whereas one who wants to identify with or impress the native speaker will tailor his/her speech to that of the interlocutor. Bell (1984), arguing from the

perspective of the Speech Accommodation Theory, therefore claims that the ultimate cause of variation in IL is the interlocutor, and not attention to speech. He argues that attention to speech is, at best, an intermediary and not an explanatory factor.

Bell's concept of variability obviously excludes task variability, for it is difficult to see how the interlocutor effect can trigger differences in styles between, for example, word list reading and sentence combining task (Tarone, 1988). We presume that, ideally, a learner who wishes to impress the interlocutor will maintain a high level of attention between both tasks, and will therefore achieve an equivalent level of targetness in the tasks. In the same vein, a learner who does not wish to identify with the interlocutor will maintain a low level of attention between tasks, with a resulting low level of targetness. The point we are making is that, if the learner were highly affected by the interlocutor then there should be no variability (or very little of it) across tasks, given that the same level of performance would be maintained in every task. This is not to say that the interlocutor effect always accounts for invariable performance. For example, the L2 learner who produces English /r/ and /l/ accurately across a range of tasks cannot be said to have a high degree of empathy towards native speakers of English. Interlocutor effect, like attention to speech cannot be subjected to empirical measurement.

Interestingly, Tarone (1988) does not reject the argument that attention to speech is not the end cause to variation. This is a departure from her earlier proposal that attention to speech is the sole cause of variability (see Tarone, 1979, 1980, 1983, 1984). Tarone (1988) argues that even if a psychological process called "attention" is postulated, we are still left with specifying features of the communicative situation which directly cause variation, or cause increased attention which in turn causes variation (1988). Like Bell she assigns to attention an intermediary function, but rejects the proposal that the ultimate end cause of variation is always the interlocutor. Tarone recognizes that task variables, such as the amount of time allowed on a task, which have nothing to do with the interlocutor, can cause variation.

3.8 Summary and conclusion

We have considered some theories and models of phonological development. The strengths, weaknesses as well as the extent of applicability of these theories and models to L2 data have been discussed. With regard to the structuralist theory, we have observed that while the concept of a universal order of acquisition and implicational markedness relations constitute important factors in L2 phonological development, the theory deals only with the surface constraints of language. This is not to reject the notion of segmental acquisition, but as

we have seen, phonological development involves much more than the pairing of segments.

The mentalistic approach of generative phonology provides a more sophisticated view of language and its acquisition. Universal Grammar consists of principles and parameters which are inherent in the human mind. Language acquisition therefore entails learning how UG principles apply to a particular language and the values that are appropriate for each parameter. Core (unmarked) parameters are distinguished from peripheral (marked) parameters, the former being set with the aid of UG, while the latter are language particular. With regard to the role of UG in L2, the conclusion is that L2 acquisition may take place with or without the UG.

Natural phonology adopts an equally mentalist appproach, but denies the existence of a special language faculty (UG) as the pivotal factor in acquisition. Rather, acquisition involves the gradual suppression of "natural" (i.e. phonetically motivated) processes in accommodating to the phonological structure of a particular language. However, we have observed that in the L2 domain, the theory fails to distinguish between suppressed transfer and developmental processes, a distinction which is vital to the understanding of L2 acquisition. Despite its limitations, it would seem that natural phonology offers a more promising explanatory framework for L2 phenomena.

The concept of markedness has been discussed as a theory in its own right, with the aim of providing a more detailed account of the phenomenon than was done within the structuralist framework. Consequently, the notion of universal implications has been explored within Eckman's Markedness Differential Hypothesis. We have argued that the L1-L2 markedness relations posited by Eckman do not account sufficiently for L2 difficulties. This inadequacy can be remedied by considering other factors such as the sonority relations holding between phonological structures.

The final model to be examined - the Labovian model - operates predominantly from a sociolinguistic perspective. It aims to account for variability in L2 acquisition on the basis of a single factor: attention to speech. The argument, however, is that variability derives from a combination of factors, and that attention to speech is at best an intermediary factor. The Labovian model, unlike the other frameworks discussed in this chapter, is essentially atheoretical in its phonological orientation: it makes no phonological claims about its findings (James, 1989). This is hardly surprising given that no theory of phonology currently available can do justice to the facts of phonological change and variation (Ferguson, 1977). The closest that has come to achieving this aim is the theory of natural phonology which attributes variation to the fortition and lenition processes operating in different speech situations.

Generative theory, as we have already noted, considers variation to be a part of performance and not competence, and therefore does not account for the phenomenon within its framework. However, variation is part of the competence of any speaker and should be accounted for by any theory of phonology.

The next chapter deals with the formulation of hypotheses which are borne out of the issues discussed in this and the preceding chapters.

CHAPTER FOUR: FORMULATION OF HYPOTHESES

4.1 Background to predictions

In Chapter One we stated the main aim of the present study which is the investigation of the processes that operate in the acquisition of liquid and final stop segments in L2 English by native speakers of Anaang. The literature has shown that three processes influence IL phonological development - namely native language transfer, developmental and universal processes of various kinds (Tarone, 1978; 1980). Our study aims to examine the extent to which phonological processes account for deviations and shortcomings in the L2 acquisition of the said speech sounds. Moreso, processes do not operate to the same degree for all sound categories. The other aim of our study is to find out which processes predominate for liquids and which for stops.

The study also focuses on the investigation of the nature of relationship between production and perception of L2 categories. The literature reviewed in Chapter Two (see section 2.6) point to some relationship between the two phenomena, though this relationship is not necessarily isomorphous. The aim is therefore to find out the degree to which difficulties at one level are reflected (or not reflected) at the other. The comparison of production and perception is restricted to liquid

speech sounds. Stops are investigated from the perspective of production alone.

The acquisition of liquids in English has been observed to pose a major problem for many Anaang speakers who find it difficult to contrast /r/ and /l/ phonologically, and phonetically in their appropriate contexts. Failure to establish such contrasts results in phonetic deviations and more frequently phonemic substitutions, the latter sometimes causing communication problems (see section 6.3 for a comparison of liquid speech sounds in English and Anaang). In summary, the learning problem for Anaang speakers consists in (a) rendering /r/ and /l/ as separate English phonemes, and (b) maintaining the phonetic differences between liquid speech sounds.

With regard to the acquisition of stop segments, the main problem lies in the realization of such stops word finally, both in citation and in connected speech (see also section 6.3. for a comparison of stops in Anaang and English). Prepausal stops remain predominantly unreleased such that it becomes difficult to tell whether they are voiced or voiceless. They may be deleted, epenthesised, devoiced or weakened to voiced homorganic fricatives, taps or approximants. While such production patterns do not necessarily impede communication, they do contribute to the "foreignness" of the learner's L2 phonological system.

The effect of L1 transfer on L2 acquisition may therefore manifest in different forms as phonetic deviations and phonemic substitutions (e.g. flapping, lateralization, stop checking and weakening), as well as perceptual confusions which result in the neutralization of phonological contrasts. IL patterns are attributed to developmental processes (a) if they represent patterns observed in the speech of children learning English as a native language (e.g devoicing, deletion, epenthesis, reduction, gliding), or (b) if they represent the approximation or overgeneralization of L2 forms, i.e. processes which are neither due to L1 transfer nor resemble those occurring in the L1 acquisition of English.

Universal factors refer to constraints due to markedness relations between phonological structures (including syllables), and L1-L2 markedness differential, where markedness is defined according to the criteria identified in sections 3.2 and 3.6, e.g implication, order of acquisition, frequency and sonority. Universal factors may also manifest as fortition and lenition processes which are related to speech styles.

A certain degree of interdependence is assumed between factors. For example, the structure of the native language may determine which developmental and universal processes occur and in what proportion (Hodne, 1985; Sato, 1984; Tarone, 1980; see sections 2.3 and 2.4 above). Furthermore a particular process may be both

transfer and developmental, transfer and universal or developmental and universal. On the whole phonological processes interrelate in different ways, both synchronically and diachronically, to shape the IL system.

In summary therefore, the present study has two aims: (a) to investigate the extent to which phonological processes operate in the L2 acquisition of stops and liquids, and (b) to examine the nature of the relationship between production and perception. In order to address these issues the following variables are taken into consideration: language proficiency (Major, 1986a, 1987a), linguistic contexts of phonological segments (Dickerson and Dickerson, 1977), sociolinguistic and stylistic determinants, i.e. communication tasks or speech styles (Edge, 1991; Major, 1986a; Tarone, 1983; Weinberger, 1987). We have noted in the preceding chapters that L2 phonological acquisition may be constrained or mediated upon by these factors.

The following procedures will therefore be adopted in the study. We shall compare performance among Anaang learners of English ranging in proficiency from low to advanced. Performance will also be compared across a range of phonetic environments varying in levels of complexity, i.e. initial, medial, final and cluster. In the case of stop acquisition only the final position will be taken into consideration. The reason is that deviations in the pronunciation of English stops by

Anaang speakers tend to be very pronounced in this position.1

We shall also compare performance between tasks differing in degrees of formality from casual to formal - namely conversation, text, sentence, word. The production and perception of liquid sounds will be compared across levels of proficiency and phonetic environments, and will be restricted to the word task. The above issues will be addressed by a quantitative analysis of performance on liquid and stop speech sounds. The basic approach to the study will be that of contrastive analysis which entails the comparison of the sound systems of Anaang (L1) and English (L2) with particular reference to the structures in question (see Chapter Six).

4.2. Research hypotheses

Given the issues raised above, the following null and alternative hypotheses are proposed.

4.2.1 Liquid consonants: production

$H_0$1a: There is no significant difference in the production of /r/ and /l/ between L2 learners at different proficiency levels.

$H_1$1a: There is a significant difference in the production of /r/ and/l/ as a function of language proficiency: L2

1 This is the investigator's own personal observation. Moreso, anecdoctal reports by teachers of English and other speakers of Anaang interviewed informally during fieldwork seem to confirm this observation.

learners at lower proficiency levels will produce more non-target variants and phonemic substitutions than learners at higher proficiency levels.

If this hypothesis is supported it would indicate that L2 learners at lower proficiency levels exhibit a greater effect of phonological processes (PP) in their IL performance than learners at higher proficiency levels. (Beebe, 1987b; Major, 1987c; Taylor, 1975). "L2 learners" refers to native speakers of Anaang learning English as a second language. "Phonological processes" is as defined in section 4.1 above. "Production of /r/ and /l/" refers to (a) the phonetic realizations of liquid sounds resulting in variable performance, and (b) the ability to distinguish liquids as separate phonemic entities. "Non target variants" refers to the non L2 phonetic realizations of /r/ and /l/. "Phonemic substitutions" refers to the replacement of one phoneme entity /r/ with another entity /l/, and vice versa.

For the purpose of this study "proficiency" is determined by educational attainment based on the number of years of exposure to L2 English in formal classroom settings. As such, four levels of proficiency are identified: low, lower intermediate, upper intermediate and advanced. A detailed description of these groups is given in section 7.2.2.2 below. "Significant difference" means a standard 95% significance level ($p < 0.05$). The operationalization of variables as stated here applies to the rest of the hypotheses which follow.

$H_0$1b: There is no significant difference in the way /r/ and /l/ are produced in different word positions in L2 English.

$H_1$1b: The production of /r/ and /l/ in L2 English varies significantly with word position: cluster and intervocalic positions will yield more non target variants and phonemic substitutions than initial and final positions.

This prediction is consistent with findings from other studies (e.g. Mochizuki, 1981). If the prediction is supported in the present study it would demonstrate a greater effect of phonological processes in the more difficult environments.

$H_0$1c: There is no significant difference in the production of /r/ and /l/ in different communicative tasks.

$H_1$1c: There is a significant difference in the production of /r/ and /l/ according to task formality: tasks eliciting less formal speech styles will yield more non target variants and phonemic substitutions than those eliciting more formal speech styles.

"Tasks" in this and other hypotheses refers to citation (i.e. word), sentence reading, text reading and conversation. They represent a continuum of styles differing in levels of formality, the least and most formal being the citation and conversation tasks

respectively. "Formality" on its part is defined according to the criterion of attention to speech, i.e. the higher the degree of attention paid to speech the more formal the style. If supported the above hypothesis would demonstrate a greater effect of processes as styles become less formal (Dickerson and Dickerson, 1977).

With regard to the three hypotheses stated above, we expect that native language factors will exert a greater influence on IL performance than developmental or universal factors. Research findings in earlier studies (e.g. Hodne, 1895; Tarone, 1980) have shown this to be the case.

4.2.2 Liquid consonants: perception

$H_0$2a: L2 learners at different proficiency levels will not differ significantly in their ability to perceive /r/ and /l/ contrasts.

$H_1$2a: There is a significant difference in the perception of /r/ and /l/ contrasts as a function of language proficiency: learners at lower proficiency levels will identify the contrasts less accurately than learners at higher proficiency levels.

"Perception" refers to the ability to identify /r/ and /l/ contrasts by assigning a phonemic label to the sounds in their appropriate contexts. If this hypothesis is supported it would demonstrate a lesser effect of processes on performance as proficiency increases.

$H_0$2b: There is no significant difference in the way /r/ and /l/ are perceived in different word positions in L2 English.

$H_1$2b: The perception of /r/ and /l/ in L2 English varies significantly with word position: cluster and final positions will be more difficult than initial and intervocalic positions.

This is consistent with findings from the study by Henly and Sheldon (1982). If this hypothesis is supported it would indicate that the perception of liquid contrasts is more difficult in some positions than in others, and therefore that the effect of processes varies with the phonological environments of the segments.

With regard to the above hypotheses we recognize the difficulty of establishing "developmental processes" for perception in the sense defined in the study. However, we can examine perception developmentally in terms of improvements that occur over time. Thus, for perception, we are concerned with transfer and universal factors.

4.2.3 Production and perception

$H_0$3: There is no significant difference in the production and perception of /r/-/l/ contrasts in L2 English.

$H_1$3: The production of /r/-/l/ contrasts in L2 English will be significantly higher than the perception of the contrasts.

"Production" in this hypothesis refers to the ability to distinguish /r/ and /l/ segments in pronunciation as separate phonemic entities. If the hypothesis is supported it would indicate that the production of liquid speech sounds has a lead over the perception of the same sounds (Flege and Eefting, 1987; Gass, 1984; Goto, 1971; Mochizuki, 1981; Sheldon and Strange, 1982).

4.2.4 Final stop consonants

$H_0$4a: L2 learners at different proficiency levels will not differ significantly in their production of word final stop consonants.

$H_1$4a: L2 learners at lower proficiency levels will exhibit more non target production of word final stops than learners at higher proficiency levels.

"Production" here and in the next hypothesis refers to the phonetic realization of word final stops. "Non target production" refers to non L2 realizations of the stops. "Final stops" in turn include the following speech sounds: /p, b, t, d, k, g/. If the hypothesis is supported it would demonstrate a greater effect of processes at lower proficiency than at higher proficiency levels, and therefore an improvement in performance as proficiency increases.

$H_0$4b: There is no significant difference in the production of word final stop consonants in different communication tasks.

$H_1$4b: There is a significant difference in the production of word final stop consonants according to task formality: tasks eliciting less formal speech styles will yield more non target production than those eliciting more formal speech styles.

As in other hypotheses this prediction is indicative of an increasing effect of phonological processes as tasks become less formal.

4.3 Summary and conclusion

The present chapter has dealt with the predictions of the study which follow from the issues discussed in earlier chapters. These predictions have been formulated bearing in mind the aims of the study which are the investigation of processes shaping IL phonology and the relationship between production and perception. The procedures for addressing these issues have been highlighted, one of which is a contrastive analysis of Anaang and English. However, before we undertake such an analysis, we must consider the sociolinguistic setting for the acquisition of L2 English. This forms the focus of the next chapter.

CHAPTER FIVE: THE LANGUAGE SITUATION IN NIGERIA

5.1 Introduction

The investigation of the acquisition of a second language is incomplete without a consideration of the setting or context in which such acquisition takes place. This chapter presents a general description of the language situation in Nigeria. Three types of language categories are identified: (a) indigenous or Nigerian languages, (b) exogenous or foreign languages, and (c) a neutral language, namely Anglo-Nigerian pidgin (see Akinnaso, 1991). Of interest to the present study, however, are the first two categories. With regard to the second we shall focus our attention mainly on English. Other exogenous languages such as French and Arabic are not discussed because they do not serve the purpose of this study.

5.2 The Nigerian languages

5.2.1 General considerations

Nigeria has approximately 400 indigenous languages1 spoken by a population of 88.5 million

---

1 There is no consensus as to the exact number of languages. It is claimed that they may be as few as 250 (Ikegulu, 1989) or as many as 500 (Odumuh, 1990). One thing worth bearing in mind, however, is that some of the languages may be dialects in reality.

(Adeniran, 1990; Bamgbose, 1971; Jibril, 1982; Omodiaogbe, 1992; *West Africa*, 30th March to 5th April, 1992), and spread across three broad language families - Niger-Kordofonian, Afro-Asiatic and Nilo-Saharan (Greenberg, 1963; Hansford, Bendor-Samuel and Standford, 1976). About 70% of the languages derive from the Niger-Kordofonian stock, a majority of which are concentrated in the south and some parts of north eastern Nigeria. The remaining 30% derive from the other two families and are found in the northern parts of the country.

The languages are further subdivided into three groups - national, regional and local - depending on a number of factors which include the demographic and geographic spread of their users, range of functions, the extent and use in formal education and the degree of official recognition (Akinnaso, 1991:32). The national or major languages are Hausa, Igbo and Yoruba according to the Constitution of the Federal Republic of Nigeria (1979). Speakers of these languages constitute about 53% of the population. Hausa is spoken by approximately 20.6 million people2 in the northern states of Adamawa, Bauchi, Borno, Jigawa, Kaduna, Kano, Katsina, Kebbi,

---

2 Available demographic figures for this and other indigenous linguistic groups are based on UNESCO's (1988) projections of Nigeria's population, i.e. prior to the 1991 census. Given the current census figure of 88.5 million as opposed to the estimated 120 million, the population of the different linguistic groups may be smaller than originally thought. However, the UNESCO estimates are still useful in that they give an idea of the size of groups in relation to one another. Hausa, for example, constitutes the largest linguistic group in the country.

Niger, Sokoto, Taraba, Yobe and some parts of Kwara state. Yoruba is spoken by about 19.9 million people in the west, in Lagos, Ogun, Ondo, Osun, Oyo and some parts of Kwara state. Igbo speakers number about 16.4 million and occupy the eastern states of Abia, Anambra, Enugu, Imo and Ikwerre in Rivers state.

Regional languages include Edo (Edo state), Efik (Cross River state), Fulfude (most northern states), Gwari (Kaduna state, FCT3), Ibibio (Akwa Ibom state), Idoma, Igala and Tiv (Benue state), Ijo (Rivers and Delta states), Igbirra (Kogi state), Kanuri (Borno and Gongola states) and Nupe (Niger state). They are spoken by 27% of the population (Akinnaso, 1991). Interestingly, there is considerable overlap between national and regional languages as the former join the latter to perform regional functions. The map of Nigeria below shows the approximate locations of these languages.

The local or "minority" languages are not represented on the map (Figure 5.1). Ironically, they constitute 95% (i.e. over 380) of the estimated number of languages, but are spoken by only a fifth of the population. More than two-thirds are found in the north, and the remaining one-third cluster around Akwa Ibom, Cross River, Delta, Edo and Rivers state in the south.

---

3 Federal Capital Territory - Abuja.

Figure 5.1: A linguistic map of Nigeria.

Adapted from Udo (1970:5)

5.2.2 The Anaang-Efik-Ibibio cluster

The last section has provided a panoramic view of the various groups of Nigerian languages. However, we are more interested in one group. We have referred to Efik and Ibibio as regional languages of Cross River and Akwa Ibom states respectively. It is worth noting that Efik is also spoken in parts of the latter state. The two form a cluster with a third language, namely Anaang, spoken in Akwa Ibom state. They are mutually intelligible sharing considerable phonological, structural and lexico-semantic similarities (Essien, 1986). Speakers of these languages can converse and understand one another, each person using his/her own language.4

Given the high degree of mutual intelligibility, they have sometimes been referred to as "dialects" of a parent language usually labelled "Efik" (Essien, 1970) or "Ibibio" (Essien, 1986, 1990). Anthropological or linguistic works of European scholars and missionaries

---

4 Cook (1969/1982) claims that these languages are not mutually intelligible by nature. He argues that speakers of Anaang can converse with Efiks not because their languages are mutually intelligible, but because they have learnt to speak Efik as an L2. He adds that some of the similarities between Efik and Ibibio are due to the influence of the former on the latter through its long use as an L2. These claims have no basis whatsoever, and nothing could be farther from the truth. Anaang speakers do not have to learn Efik to understand it. The same goes for Efik speakers who understand Anaang perfectly well without learning it. In the same vein Ibibio speakers understand and are understood by speakers of the other languages. It is surprising that Cook claims some degree of intelligibility between Efik and Ibibio based on the use of the former as an L2, but does not do the same for Anaang and Efik even though speakers of Anaang have used Efik as an L2.

are often evoked as justification for the classification.5

However, there is some controversy over the description of Anaang, Efik and Ibibio as "dialects", and over the labelling of the parent language. While we do not wish to be part of this controversy, suffice it to say that the arguments put forward often stem from considerations other than linguistic. Referring to the Anaang-Efik-Ibibio situation, Connell (1991:9) rightly argues that "the actual deciding factor as to whether languages should be classified as dialects or languages depends on political boundaries or cultural identities of the groups involved." Generally, speakers of these languages believe that they belong to distinct, but related ethno-linguistic groups. As such they see themselves as speaking "related languages" rather than as speaking "dialects" of one language.6 For most people the term "dialect" implies "inferiority" or "subordination" to another linguistic group. The desire for each group to assert or create a separate identity for itself is, therefore, quite strong. With regard to

---

5 See for example, Forde and Jones (1967), Jeffreys (1935), and Westermann and Bryan (1952).

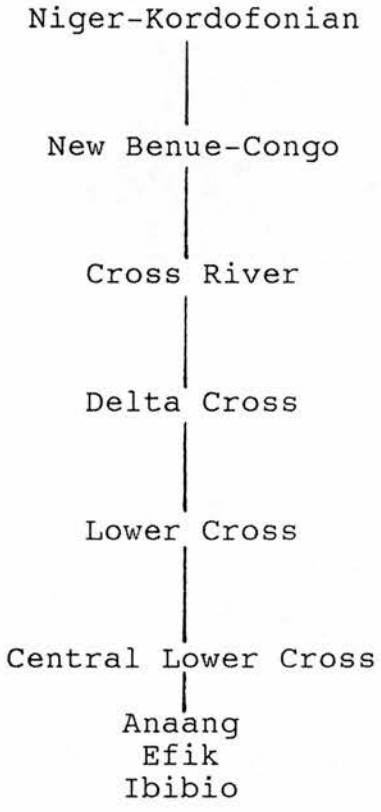
6 Nevertheless, the Ibibios tend to regard the Anaangs and the Efiks as part of the main Ibibio stock, speaking variations of Ibibio. They therefore see them as Ibibio people (cf. Abasiattai, 1991; Udo, 1983). This has been resented by many Anaang and Efik speakers who maintain that they are not Ibibios, though they are related to them (cf. Udo, 1983). Some Efiks have gone as far as denying any links with the other groups. This is more of a political than a linguistic issue.

the present study, Anaang, Efik and Ibibio are referred to as languages rather than as dialects, moreso because they have clearly discernible dialects of their own. However, the main focus of our attention is the Anaang language.

5.2.2.1 Genetic classification

According to Greenberg (1963), the Anaang-Efik-Ibibio cluster form part of the Benue-Congo branch of the Niger-Congo language family, which in turn belongs to the higher Niger-Kordofonian family. A further sub-classification groups these languages under the Cross River 2 sub-branch of the Benue-Congo. Greenberg's classification has since undergone modifications. A recent model proposed by Williamson (1989, cited in Connell, 1991), for example, places the three languages in the Lower Cross sub-group of the Delta Cross branch of Cross River. The latter in turn belongs to the enlarged New Benue-Congo. Following on from there, Connell (1991) establishes a separate sub-group of the Lower Cross called Central Lower Cross, under which he places Anaang, Efik and Ibibio. These genetic affiliations are shown in Figure 5.2 below.

Figure 5.2: Genetic classification of Anaang-Efik-Ibibio.



5.2.2.2 Geographic and demographic distribution

Speakers of the three languages live east of the Niger Delta region of Nigeria. The Efik areas border the Republic of Cameroun to the east, the Ejagham and Ibibio speaking areas to the north and west respectively. The Ibibios are sandwiched between the Efiks and the Anaangs (their most westerly neighbours). They are also bounded to the north by the Igbos and to the south west by the Obolos who are straddled between Akwa Ibom and Rivers states. The Anaang areas are bordered to the east and

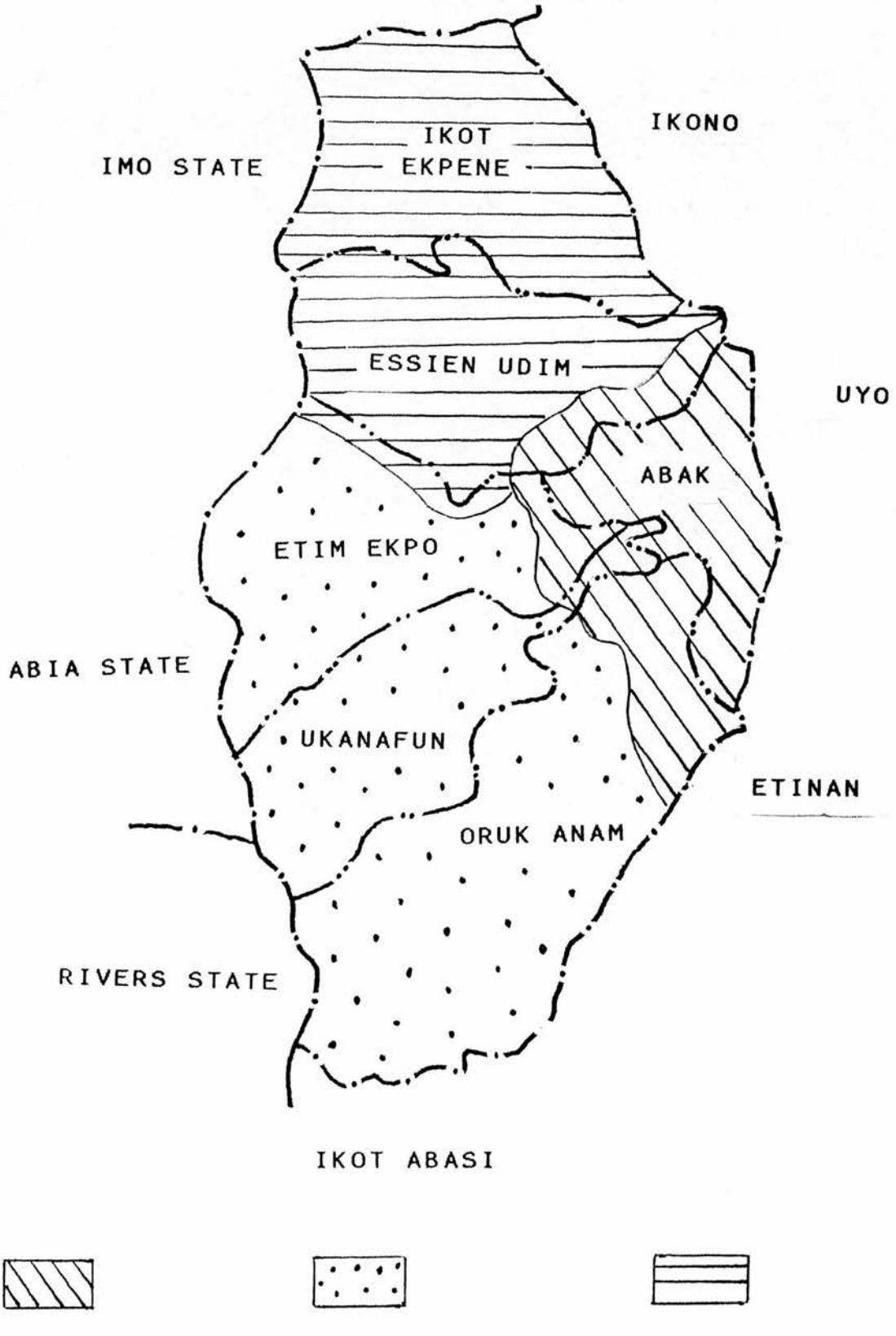
south by the Ibibios, and to the north and west by the Igbos.

While Efik remains the best known language of the group by virtue of the early contact of Efik people with European traders and missionaries, and by virtue of having enjoyed the status of a *lingua franca* for over a century (Westermann and Bryan, 1952), it is by no means the largest group. Less than 500,000 people speak the language as a mother tongue (Connell, 1991). The largest language of the group is unquestionably Ibibio, having an estimated 3.5 million speakers.7. It ranks fourth after the major languages Hausa, Igbo and Yoruba (Essien, 1990). Anaang constitutes the second largest group of the cluster, and is spoken by about 1 million people.

For the purpose of this study Anaang speaking areas are designated broadly as "Ikot Ekpene", "Abak" and "Ukanafun", and the three major dialects of the language also go by these names. Alternatively, we could classify the dialects as northern (Ikot Ekpene) eastern/central (Abak), and western/southern (Ukanafun). This classification is highly subjective, but it gives us some idea of the geographical location of the dialects as shown on the map in Figure 5.3 below. It is not based on current political boundaries, otherwise known as Local

7 This figure is actually higher than that of the entire population of Akwa Ibom state (2.36 million according to the 1991 census), where Ibibio and other languages are spoken. Again, the estimated figure given above was based on UNESCO projections.

Figure 5.3: Map showing the location of major dialects of Anaang.



Government Areas (LGA), since the dialects cut across such boundaries. For example, the Ukanafun dialect is spoken in Ukanafun and Oruk Anam LGAs as well as in most parts of Etim Ekpo LGA. The Abak dialect is spoken in Abak and those parts of Ukanafun, Etim Ekpo and Oruk Anam LGAs which share the same boundaries with Abak LGA. Thus, there is considerable overlapping between dialects along political boundaries. The dialects differ mainly in the phonological domain with regard to certain phonemic/allophonic alternations, and intonation (Essien, 1970/71).

5.2.2.3 Previous studies

Of the three languages Efik has enjoyed a long tradition of distinguished linguistic scholarship. This is evident in the number of publications that have emerged since the arrival of the Scottish Presbyterian missionaries in Calabar in 1846 (Cook, 1990). Scholars who have written on Efik include H. M. Waddell (*A Vocabulary of Efik or Old Calabar language*, 1849), Rev. H. Goldie (*Dictionary of Efik Language*, 1862; *Principles of Efik Grammar*, 1868; *Efik Grammar in Efik*, 1874; *Efik Grammar in English*, 1874), Ida Ward (*The Phonetic and Tonal Structure of Efik*, 1933), W. Welmers (*Efik*, 1966), T.L. Cook (*The Pronunciation of Efik for Speakers of English*, 1969/82). Other works have included translations of the Bible and other religious books from English into Efik. The list is by no means exhaustive. Through the

work of these scholars and others, Efik has exerted considerable though little recognized influence on African linguistics and general linguistic theory (Connell, 1991; Cook, 1990).

In comparison, the linguistic study of Ibibio language is a fairly recent phenomenon. As such, relatively fewer materials are available on the language. One of the earliest works on Ibibio dates back to 1935 and was undertaken by M. D. W. Jeffreys (*Old Calabar and Notes on the Ibibio Language*). An article entitled "Ibibio Verb Morphology" was published by D. C. Simmons in 1957. Other works include E. M. Kaufmann's *Ibibio Grammar* (1968) and *Ibibio Dictionary* (1985), O.E. Essien's *The orthography of Ibibio Language* (1983a), *A Grammar of Ibibio Language* (1990), and a number of articles (see Essien, 1983b, 1983c, 1985a, 1985b). Other contributions to the study of Ibibio have been made by Connell (1989, 1991), Edet (1990) and Urua (1990).

Anaang, unfortunately, has not enjoyed the status of a written language as have Efik and Ibibio. There are only a handful of materials available on the language. The first attempt at compiling a word list in Anaang was undertaken by S. W. Koelle in *Polyglotta Africana* (1854). His informant was an ex-slave who had returned from America to settle in Sierra Leone. Despite his long separation from Anaang communities, the informant was able to provide as near accurate data as possible. From the time of Koelle's publication to the early 70's no

substantial work was carried out on the language. In 1970/71 two articles were published by U. E. Essien. The first, entitled "The vowel structure of Efik" is the first generative analysis of Efik and Anaang vowel systems. The second consists of a review of Koelle's *Polyglatta Africana*, in which Essien compares Koelle's transcriptions with present day Anaang.

However, the past two years have witnessed a resurgence of interest in the linguistic study of Anaang. As a first step, the newly formed Anaang Language Writers Association has undertaken to produce an acceptable orthography and dictionary of the language. While our study focuses primarily on language acquisition, we hope that it will contribute in some way to the on-going study of Anaang.

5.2.3 Nigerian languages in education

The use of Nigerian languages in education dates back to the era of the early missionaries. In order to ensure the effective propagation of Christianity and the internalization of its concepts, the missionaries taught the local people to read the Bible and other religious documents in their native languages (Akinnaso, 1991; Odusina, 1989). As Fafunwa (1974) points out, they "discovered that a child or man learns better, absorbs more and appreciates better in his mother tongue" (p. 90). This was consistent with their overall goal at that time, which was to train local clergymen for the church.

The colonial administration was to discourage mother tongue education through the enactment of an Educational Ordinance (1882) which accorded priority to English education. However, a change in policy was effected when in 1920 the Phelps-Stokes Commission recommended the use of Nigerian languages as media of instruction in the lower primary school, while retaining English in the upper primary and secondary schools (Akinnaso, 1991:39).

Subsequent policies on language education upheld this recommendation. However, in 1977 when the new National Policy on Education (NPE) was promulgated, certain changes were introduced. In addition to advocating the use of local languages in pre-primary education (section 2, paragraph 11) and in the first three years of primary education (section 3, paragraph 15/4), the new NPE requires that each child learns one of the three major languages apart from his mother tongue (section 1, paragraph 8) (Federal Ministry of Information, 1977).

The curricula of primary and secondary schools have therefore been modified to reflect the multilingual orientation of the policy. In the last three years of primary schooling when the local language ceases to serve as the medium of instruction, it is taught as a subject. In the Junior Secondary Schools two Nigerian languages - the mother tongue and one national language - are learnt. Only the national language is retained in the Senior

Secondary School curriculum. At the university level the following languages are approved for teaching: Edo, Efik, Hausa, Ibibio, Igbo, Kanuri and Yoruba (Essien, 1990).

In order to achieve the policy's objectives, both the federal and state governments have encouraged the development of orthographies for more Nigerian languages, and the production of text books in these languages. The famous "Rivers Readers Project" sponsored by the Rivers State government in conjunction with the University of Port Harcourt has produced primers and supporting materials for all the languages and major dialects of the state (Williamson, 1980). Other projects have been undertaken by individual language boards, e.g the Ibibio Language Writers Association. Language centres have also been set up in many parts of the country for the purpose of enhancing the study of Nigerian languages.

Nevertheless, it is doubtful if all the objectives of the NPE are pursued with much vigour. For example, the existence of multi-national pre-primary and primary schools implies that the implementation of mother tongue education is not possible in such schools. On the contrary, most rural schools continue to use the local languages beyond the stipulated three year period. Obstacles to the effective implementation of the policy include (a) the lack of adequate personnel, material and financial resources to support the teaching of the three major languages in particular, (b) the unwillingness by state governments, as a result of financial constraints,

to sponsor the teaching of any Nigerian language which is not native to their states, (c) the absence of a federal machinery to monitor the implementation of the policy in the country, and (d) the lack of motivation on the part of the learners who see no need for learning languages which have no bearing on their future careers (Akinnaso, 1991; Banjo, 1989; Emenyonu, 1989). Perhaps the greatest obstacle is the prestige that English continues to enjoy to the detriment of the local languages.

5.3 The English Language in Nigeria

5.3.1 Historical background

The first Europeans to visit what is now Nigeria were the Portuguese traders who called at the ancient city of Benin in the 15th century (Awonusi, 1990; Ayandele, 1966; Fafunwa, 1974). From this contact emerged a form of pidgin Portuguese which became the trade language of coastal Nigeria at that time. Relics of its use can still be found in present day Anglo-Nigerian pidgin in words like "sabi" (to know) and "palaver" (quarrel). Portuguese influence declined following the arrival of British traders and the subsequent development of "broken" and later pidgin English as the legitimate language of trade. By the 18th century English was already in use by Nigerian traders, rulers and bureaucrats (Awonusi, 1990).

Therefore, when the missionaries arrived in the 19th century, following the abolition of the Atlantic slave trade, they found people who were keeping not only regular accounts, but also diaries in some form of English, and had been doing so for a century or more (Ajayi, 1965). In Old Calabar certain Efik kings - "Antera" Duke, Eyo Honesty II and Eyamba V - are reported to have kept diaries of their governance and trade transactions with the Europeans in English (Ayandele, 1966; Awonusi, 1990). When Rev. Hope Waddell arrived in Calabar in 1846 he found records and journals dating 1767 and earlier (Fafunwa, 1974).

In the previous section it was remarked that the missionaries initially favoured vernacular education but were discouraged from doing so by the colonial government. However, prior to the enactment of the 1882 Education Ordinance by the government there was considerable demand for English education by the local people themselves. In Calabar, for example, where trading was the main occupation, the people were more interested in education than in Christian evangelism. They had their own religion and therefore did not need a new one. All they wanted was to acquire English for trading purposes and the learning of trade skills.

The missionaries not only succumbed to these and other pressures, but realized, in the process, the need for the formal training of the local clergy in the language of the new religion - English. Schools sprang

up in different parts of southern Nigeria, the primary aim being to train catechists, mass servers and preachers. The teaching of English was initially carried out by native and non native speakers alike: British, Germans, French and liberated slaves from Sierra Leone.

The colonial period saw the institutionalization of English education in Nigeria through the 1882 Education Ordinance which made the language compulsory in mission schools. Not only was English taught as a subject, it became the medium of instruction for other subjects. The main text books for teaching English were the Bible and English classical literature books. As such the English spoken and used at that time was "characterised by the formality associated with Victorian English mediated by some style of King James biblical English" (Awonusi, 1990:32).

The emphasis on English education by the colonial administration is hardly surprising. It needed clerks, and lower cadre staff for the effective implementation of its policies. It was not interested in the religious inclination of the people as long as they could speak and write good English. Mission schools were therefore converted into clerk making machines, and the condition for obtaining government grants rested on the success of pupils in English. The axe fell on non-English European missionaries who were replaced by native English speaking teachers. Schools were set up by the government in the

Muslim north where missionaries were not allowed to operate.

English spread from the schools into the Nigerian society as the colonial administration cultivated and made it the language of all official interactions. It became the main language of communication, a means of social, educational and economic advancement. It is probably the most important legacy left by the colonial administration in terms of the status and functions it has assumed.

5.3.2 The status and functions of English in Nigeria

The removal of English from its native British environment, its implantation and development in a non native one means that it assumes a status other than that of a mother tongue. In Nigeria English plays the role of a second language where it serves a variety of functions. It is the language of government, business and commerce, the mass media, literature, education, much internal and external communication (Bamgbose, 1971:35).

English serves as the official language of government and administration: all government records, administrative instructions and minutes are all kept in English. In the political domain it is used along with indigenous languages. The Nigerian Constitution (1979), for example, requires that the business of the National Assembly be conducted in English and the three major languages (section 51). In the State Houses of Assembly,

English is to be used in conjunction with any of the state languages (section 91).

With regard to the judiciary, it remains the language of court proceedings and the documentation of legal issues. This applies to all court systems (federal, state, supreme and appeal), except the customary courts. In the latter English is used alongside local languages (Awonusi, 1990), and sometimes not at all. Business records and contracts are also written in English.

While the recent years have witnessed an increasing use of Nigerian languages in the media, English remains the dominant language in this domain. Most newspapers, journals and periodicals are published in English, and a majority of radio and television programmes are in English. However, as Awonusi (1990) points out, there seems to be greater dependence on English on television than on radio.

English also serves for much internal as well as external communication. The multilingual composition of Nigeria calls for a common language for the purpose of communication and ethnic understanding - a role which English has played in the absence of an indigenous *lingua franca*. The choice of a *lingua franca* for Nigeria is a very sensitive issue, the details of which do not serve the purpose of this study. Suffice it to say, however, that English acts as a unifying force in a delicate multilingual set up. As Odumuh (1989:15) confirms,

> it has the advantage of being a compromise candidate at least for the interim, since the adoption of one of the so-called major indigenous languages as the national language would have meant the alienation of large sections of the citizenry, especially in an atmosphere bedevilled by ethnic nationalism (or parochialisms).

While English plays this vital role, it is claimed that no more than 5% of the Nigerian population are competent users of the language (Adeniran, 1990; Bamgbose, 1971; Emenyonu, 1989). Such users are to be found among the educated elite. Unfortunately, there are no statistical data to support this claim. Externally, English constitutes the main medium of interaction between Nigeria and other countries. Perhaps the most important function of English is its use in the education system. This is examined in a separate section below.

5.3.3 English in Education

The position of English as a language of education was solidified by the provisions of the 1882 Education Ordinance, as pointed out earlier. English is now taught as a subject in the first three years of primary education during which Nigerian languages are used as media of instruction. From mid-primary to university education, it becomes the main medium of instruction while retaining its position as a taught subject (Awonusi, 1990; Bambgose, 1971; Odumuh, 1989).

However, the point at which the transition from mother tongue to English education occurs varies from place to place. We remarked in section 5.2.3 that most rural schools tend to use the mother tongue beyond the stipulated three year period. This may be attributed to the fact that by the fourth year of primary education, i.e. the official year of transition, the children have not acquired enough English to cope with English medium instruction. It must also be noted that most of the rural schools are situated in monolingual villages where the need for everyday interaction in English is minimal.

Nevertheless, extended mother tongue education is not restricted to rural schools. Many urban schools, especially those in the Yoruba speaking areas where the Ife Six Year Primary Project is being implemented, use the mother tongue throughout primary education as a matter of policy, with English being taught only as a subject. Another trend altogether is observed in some states (e.g. Edo, Delta) where Anglo-Nigerian pidgin, despite its stigmatization, is the *lingua franca* and is used as a medium of instruction in primary schools whether urban or rural (Akinnaso, 1991; Omodiaogbe, 1992). At the opposite extreme are the multi-national pre-primary and primary schools where English is the sole medium of instruction from the first year.

The consequence of such variable implementation of the language policy is that children bring different experiences to bear upon secondary education, where the

only medium of instruction is English. As such, some are more advantageously placed than others, at least initially, with regard to their proficiency in English. Nigeria operates a two-tier secondary school system, and a pass in English among other subjects is required before students are upgraded from the Junior to Senior Secondary School levels. Students are also required to obtain at least a pass grade in English in the Senior Secondary School Certificate or General Certificate of Education (G.C.E. O/A levels) examinations. Additional exposure to English is obtained during English literature classes in both the junior and senior sections of the secondary school system. At the university level the least entry requirement is a good pass in English. Moreso, a General English course is made compulsory for all first year students as a means of remedying any deficiencies in English competence.

From the above discussion it is obvious that success at each level of the educational system depends largely on competence in English, moreso because such competence is relevant to future intellectual pursuits and success in careers (Banjo, 1989; Emenyonu, 1989). Thus, while the language policy may promote the study of Nigerian languages in addition to English, it is the mastery of the latter that is emphasized for the obvious reasons stated above, and because it is the official language of the country. As such there is a growing concern over the standard of English which is said to be

declining. This has been attributed to several factors which we shall not delve into, though a lot probably has to do with the type and quality of English received at school coupled with interference effects from the different mother tongues, and the normal processes of language development.

5.3.4 Nigerian English

5.3.4.1 Definition and classification

An examination of the language situation in Nigeria would be incomplete without a consideration of the notion of Nigerian English (NE). Questions arise as to the reality of its existence, its definition and classification.

There is (or should be) no doubt as to the existence of NE given the historical development of English in Nigeria, its status as a second language and the various functions it performs in the Nigerian society. However, many educated Nigerians continue to deny the reality of NE "even though their own speech and usage provide ample evidence of its existence" (Bamgbose, 1982:99). They believe that the English they speak and write is no different from the British Standard English (BSE) or Received Pronunciation (RP). Others recognize its existence but describe it as sub-standard and inferior (Odumuh, 1990). This group also denies using NE. Still there are other Nigerians who think that "pidgin"

is what is meant by NE (Jibril, 1982). Bambgose (1982) has pointed out that most of those who refuse to accept the existence of a Nigerian English are worried about the impact such a model might have, particularly in the area of language teaching. They fear that, sooner or later, such a variety may degenerate into a different language like Anglo-Nigerian pidgin.

Nevertheless, the definition of NE is more difficult than envisaged. As a matter of fact, Odumuh (1990) notes that NE is better described than defined. In general terms NE may be defined as that form of English characterized by marked linguistic, social and cultural features of the Nigerian environment. From the linguistic perspective, it may be said to bear the marks of the different indigenous languages with which native English has come into contact, and over which it has been superimposed (Akere, 1982). However, such "indigenous" influences alone do not account for the differences that exist between NE and BSE. Other differences may be traced to normal language development processes or teaching models.

Despite the controversy over the existence or non existence of NE, linguists have classified NE into various types. Using level of education as a parameter, Brosnahan identifies the following varieties:

> Level 1: pidgin - spoken by those without any formal education;

> Level 2: spoken by those who have had primary education;
>
> Level 3: spoken by those who have had secondary secondary education and is marked by increased fluency, wider vocabulary and avoidance of level 1 usage;
>
> Level 4: spoken by those with university education and is close to standard English, but retains some features of levels 2 and 3.
>
> (cited in Bamgbose, 1982:100).

Banjo takes the definition even further and introduces variables of international intelligibility and social acceptability. He also identifies four varieties which include:

> Variety 1: Spoken by those with an imperfect knowledge of English and marked by wholesale transfer of phonological, syntactic and lexical features of the indigenous languages. Neither socially acceptable nor internationally intelligible;
>
> Variety 2: Close to standard British English in syntax, but with strongly marked phonological and lexical peculiarities. Socially acceptable but not internationally intelligible;
>
> Variety 3: Close to British English in syntax and semantics; similar in phonology but different in phonetic features and certain lexical peculiarities, socially acceptable and internationally intelligible;
>
> Variety 4: identical with British English in syntax and semantics, and having identical phonological and phonetic features of native English. Maximally internationally intelligible but socially unacceptable.
>
> (cited in Bamgbose, 1982:100/101).

There is remarkable similarity between the two classifications, except that while Brosnahan treats "pidgin" as a native variety of NE, and native English as a separate variety, Banjo excludes "pidgin" from his scheme, but treats native English as a variety of NE. However, Banjo provides a better classification by explicitly stating their characteristics. Brosnahan is mainly concerned with the speakers' level of education, and bases proficiency in English on this criterion. One problem with using education as a measure of proficiency is that there are speakers who perform below or above their expected level of competence based on educational attainment. Bamgbose (1982) has therefore suggested that proficiency should be measured on the basis of the speaker's output rather than on his/her level of education as the former does not necessarily reflect the latter. Moreso, other variables like innate ability, intelligence and the opportunity for the use of English should be taken into account. While these classifications provide a general description of NE, the present study is more concerned with its spoken form.

5.3.4.2 Nigerian Spoken English

It is obvious that the major difference between English in Nigeria and English in other countries lies in the spoken form (Bamgbose, 1971). What is referred to as Nigerian Spoken English (NSE) ranges from the heavily accented type (with marked interference from the mother

tongues at the segmental and suprasegmental levels), to the "near native" type (with RP phonological deep structure and Nigerian surface structure) (Williams, 1983). Mid-way along the continuum is the slightly accented type which has more mother tongue influence at the suprasegmental than at the segmental level.

There is as yet no comprehensive description of NSE in the sense that there is of native British Spoken English. This may be due partly to the stigmatization of Nigerian English as a whole, and partly to the fact that "British English" is still the "prescribed norm" in the education system. Another factor to consider is the existence of different regional varieties, which raises the question of the choice of a model of description. Nonetheless, this problem seems to have been partly resolved in that available descriptions are based on what is termed "educated Nigerian English" (ENE). This variety is to some extent devoid of local variation, and corresponds to Banjo's type 3 or William's near native category.

However, it is worth noting that descriptions of NSE consists of accounts of isolated segmental and suprasegmental phenomena. Two approaches are used: interference and deviation (Bambgose, 1982). The interference approach identifies features that are attributable to L1 influence, while the deviation approach compares observed Nigerian usage with native English and labels all differences as "deviant".

Deviations may be due to (a) interference, (b) the imperfect attempt to reproduce the target sounds resulting in features that are typical of most Nigerian speakers of English, and (c) to the normal process of language development involving the narrowing or extension of sounds. Some general features of NSE include:

(a) A reduced vowel system involving different substitutions. A basic seven to fourteen vowel system operates (Awonusi, 1990), though Ekong (1982) identifies sixteen vowel contrasts out of twenty-four in native English including diphthongs and triphthongs which are generally monophthongized. E.g. [e] or [ɛ] is substituted for [ei] as [mek]/[mɛk], "make".

(b) The absence of a distinction between long and short vowels as in [bi:t] "beat" and [bɪt] "bit".

(c) The general absence of the interdental fricatives /θ/ and /ð/ which are often replaced by /t/ and /d/ or /s/ and /z/ respectively. The absence of voiced palato-alveolar fricative /ʒ/ for which /s/, /z/ and /d / are substituted.

(d) Syllable-timed instead of stress-timed rhythm.

(e) The operation of processes such as cluster reduction, syllabification and vowel epenthesis.

Though these features are common to most speakers of Nigerian English, some are peculiar to certain geographical locations. For example (Jibril, 1982) points out that most southern Nigerians substitute [t] and [d] for native English /θ/ and /ʒ/, while the northerners

realize the same sounds as [s] and [z] respectively. The former constitute what he terms Type 2 speakers, and the latter Type 3 speakers. Thus, with regard to vowel substitution, Type 2 speakers replace /ʒ/ with [ɔ] as in "fur" [fɔ], with [ɛ] as in "birth" [bɛt] and [a:] or [ɛ] as in "earn" [a:n] or [ɛn]. For Type 3 speakers /ʒ/ becomes [a] so that "fur" is homophonous with "far" [fa:]. "Birth" is realized as [ba:z] and "earn" as [a:n].

While the regional factors identified above cannot be overlooked, the classification of NSE on the basis of a "north/south" dichotomy is too broad. It is not the case that all southerners or northerners speak English in the same way. The Yoruba man definitely has a different accent from his Igbo and Ibibio counterparts in the south. The Tiv speaker has a different accent from his Hausa or Kanuri counterparts in the north. As such the effect of individual mother tongues on English is quite strong.

5.4 Summary and conclusion

The main concern of this chapter has been to describe the language situation in Nigeria in order to provide a sociolinguistic background to the present study. A highly multilingual society is revealed with English playing the leading role against a background of diverse Nigerian languages. Given its long contact with these languages and the various functions it performs, English has gradually developed away from its native

form, reflecting more and more of features associated with Nigerian speakers. Yet the existence of a Nigerian variety of English is not acceptable to all, especially the educated elite. This phenomenon has implications for the present study in that it raises questions regarding the model of English acquired by the Nigerian child, and the model to be adopted in this study. This issue is considered in the next chapter which presents a description of English and Anaang.

CHAPTER SIX: LINGUISTIC STRUCTURES TO BE INVESTIGATED

6.1 Introduction

The present chapter focuses on the phonological structures that have been selected for investigation. A contrastive approach is adopted which entails, firstly, an overview of the sound systems of English and Anaang, and secondly, a detailed description of liquid and stop segments in the two languages. The problems encountered by Anaang speakers in the acquisition of English liquid and stop speech sounds were highlighted in Chapter Four (see section 4.1).

The model of English adopted for the purpose of our study is the educated British English commonly referred to as Received Pronunciation or RP. In reality this accent is used by a small proportion of the entire population of the English-speaking world. However, for historical/social reasons it is widely accepted as the standard form of pronunciation, and constitutes the norm in most language teaching and learning circles (Gimson, 1980; Knowles, 1987). This has led to its being the most comprehensively described model of English, and the one used as a basis for textbooks (Roach, 1983).

While we recognize the existence of developing forms of English, e.g. Nigerian or Ghanaian English, it is worth noting that these varieties are in themselves the product of attempts at accessing the native British

model. The problem is that in many quarters they are still not regarded, at least theoretically, as suitable models in a language learning/teaching situation. Another problem, as noted in Chapter Five, is the absence of **comprehensive** descriptions of the new models of English. Lastly, the RP is the model used in most second language research. In the present study, therefore, we intend to compare the English IL performance of Anaang speakers with this **ideal** model.

Of the three dialects of Anaang referred to in Chapter Five, the Ukanafun dialect has been chosen for descriptive purposes. References will be made to other dialects, and to related languages where the need arises.

6.2 The sound systems of English and Anaang

The sound system of a language consists of segments as well as suprasegmentals. Consonants and vowels constitute the segmental units which combine to form syllables. Superimposed on these syllables are suprasegmental features such as stress, intonation and tone. This section presents an overview of the structures which make up the sound systems of English and Anaang. The aim is to provide a general background to the more detailed description of selected phonological structures which follows.

6.2.1 English

The description of English Phonology presented below is based on Abercrombie (1967), Gimson (1980), Kreidler (1989), Ladefoged (1982), O'Connor (1973) and Roach (1983).

6.2.1.1 Consonants

English has twenty four consonant phonemes which may be classified into two broad categories according to manner of articulation:

> (i) those which involve a total closure or stricture causing friction such as stops (or plosives), fricatives and affricates, and which are associated with a noise component,
>
> (ii) those which involve only a partial closure or unimpeded oral or nasal escape of air such as the nasals, approximants (i.e. glides and liquids).
>
> (Gimson, 1980: 151).

Both categories are therefore distinguished by the feature *sonorant* expressed in binary values as [+sonorant] (i.e. nasals, liquids, glides) and [-sonorant] (i.e. stops, fricatives and affricates). Consonants in the first category are further distinguished by the presence (+) or absence (-) of voicing, though it is argued that the degree of breath and muscular effort involved in their articulation constitutes a more important distinctive feature as voicing opposition may be lost in certain situations, e.g word initially and finally. Voiceless consonants tend to

be realized with more force than their voiced counterparts, and are referred to as fortis (or strong). Voiced consonants are lenis (or weak). The second category of consonants is typically voiced, and frequently frictionless sharing many characteristics with vowels.

Figure 6.1 below presents the phonemic inventory of English. Six oral **stop** phonemes are identified according to their place of articulation - bilabial, alveolar and velar. The voiceless fortis and voiced lenis sets include /p, t, k/ and /b, d, g/ respectively. A more detailed account of stops is presented in section 6.3.3.

The **fricatives** are divided into four correlated pairs, labiodental (/f, v/), interdental (/θ, ð /), alveolar (/s, z/), palato-alveolar (/ʃ, ʒ /) on the one hand, and /h/ without a correlate, on the other. The lenis fricatives /v, ð , z, ʒ / tend to be fully voiced only when they occur between voiced sounds as in "heavy" [hɛvi], "father" [fɑ:ðə], "easy" [i:zi] and "treasure" [trɛʒə]. In initial and final positions voicing may be partial or inexistent. /ʒ/ does not occur initially in English. The fortis fricatives /f,θ , s, ʃ / as expected are [-voice]. This fortis/lenis distinction, however, does not apply to the glottal fricative /h/ which occurs in all positions except finally.

English has only two **affricate** phonemes, the palato-alveolar fortis/lenis pair /tʃ, dʒ/, which occur in all positions, e.g initially in "chain" [tʃein],

"Jane" [dʒein], medially in "richer" [ɹitʃə], "major" [meidʒə] and finally in "catch" [kætʃ], "badge" [bædʒ]. The voicing characteristics for /dʒ/ are the same as those for voiced fricatives. The voiceless fortis counterpart is usually accompanied by aspiration in

Figure 6.1: Chart of English consonant phonemes

Manner of articulation	Place of articulation: B/bl	L-dtl	Int-dtl	Alv	Pal-alv	Pal	Vel	Glo
Plos	p b			t d			k g	
Fric		f v	θ ð	s z	ʃ ʒ			h
Affric					tʃ dʒ			
Nas (plos)	m			n			ŋ	
C/tral approx	w			ɹ		j		
L/tral approx				l				

N.B: B/bl = bilabial; L-dtl = labio-dental; Int-dtl = inter-dental; Alv = alveolar; Pal-alv = palato-alveolar; Pal = palatal; Vel = velar; Glo = glottal.

initial accented syllable, but is only weakly aspirated word finally. On the contrary, /dʒ/ never undergoes aspiration.

The three **nasal** phonemes of English correspond to the oral plosive areas of articulation: bilabial /m/, alveolar /n/ and velar /ŋ/. In terms of distribution /ŋ/

differs from other nasals: it occurs only medially and finally, while /m/ and /n/ are equally distributed across environments. Like vowels, /n/ and less commonly /m/ and /ŋ/ perform syllabic functions (see 6.2.1.3 below).

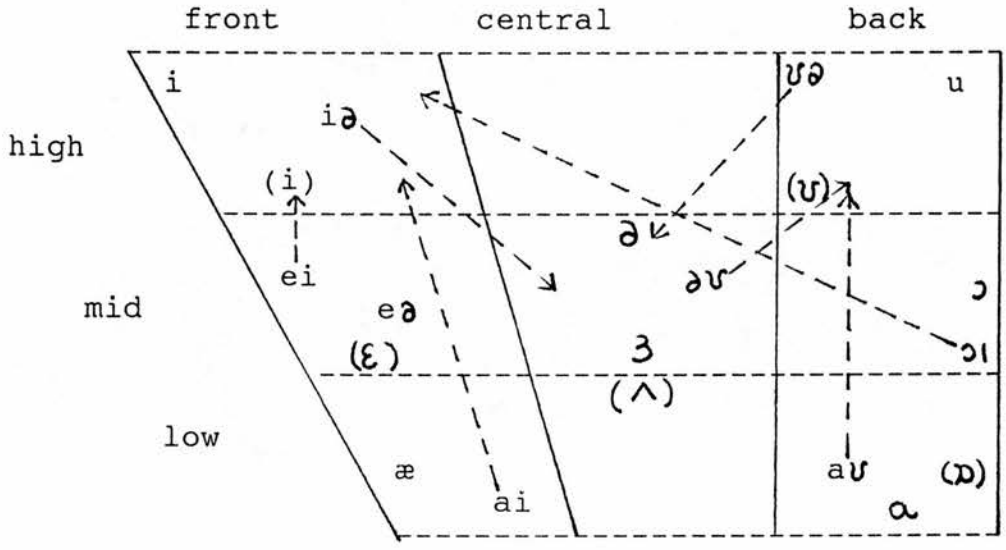
The **approximant** phonemes have neither the closure nor the friction components characteristic of other consonants. They consist of (i) the **glides** /w, j/ which are phonetically like the vowels /i:, u:/, but phonologically like consonants - they occur initially before vowels or in an initial cluster preceding a syllabic sound, e.g. "wet" [wɛt], "yet" [jɛt], "twin" [twin], "pure" [pjʊə]: and (ii) **liquids** which include the alveolar lateral /l/ [+lateral] and the post alveolar approximant /r/ [-lateral]. The lateral differs from the rest of the approximants in that the airstream passes through the sides of the tongue instead of the centre. Liquids are treated in greater detail in section 6.3.1 below.

6.2.1.2 Vowels

English has twenty vocalic phonemes which are classified according to their quality, length, complexity and tenseness. Vowel quality is determined by the shape of the oral cavity which in turn depends on the following features: **tongue height** (high, mid, low), **tongue advancement** (front, central, back) and **lip shape** (round, unrounded) (Kreidler, 1989). These features are represented in the vowel chart below.

The chart shows eight front vowels of which three are high /i:, i, iə/ as in "beat" [bi:t], "bit" [bit], "fear" [fiə], three mid /ɛ, ei eə/ as in "set" [sɛt], "day" [dei], "pair" [peə], and two low (/æ, ai) as in "mat" [mæt], "tie" [tai]. These vowels are marked by the absence of the feature *round*, and may therefore be

Figure 6.2: Chart of English vowels.



N.B: Short vowels are shown in parentheses. Dotted arrows indicate the gliding position of diphthongs.

represented as [-round]. The central vowels consist mainly of the mid vowels /ə, ɜ, əʊ/ as in "about" [əbaut], "bird" [bɜ:d], "load" [ləʊd], and the low vowel /ʌ/ as in "but" [bʌt]. Like their front counterparts they lack the feature *round*. Seven back rounded [+ round] vowels are

distinguished with /u:,ʊ ,ʊə ,/ occurring as high vowels in "food" [fu:d], "put" [pʊt], "tour" [tʊə]. The mid back vowels include /ɔ:, ɔi/ as in "board" [bɔ:d] and "ploy" [plɔi], and the low vowels /ɒ , ɑʊ/ as in "gone" [gɒn] and "house" [haʊs]. The last vowel in this category is the low back /ɑ/ which differs from the rest by the absence of the feature *round*.

With regard to length vowels may be classified as long or short, the latter consisting only of pure vowels (or monophthongs), and the former of some pure as well as complex vowels (diphthongs). On the chart short vowels are parenthesized to distinguish them from the two categories of long vowels. The effect of complexity in diphthongs is produced by the presence of two vowels, one gliding into the other. The first part of the diphthong is usually much longer and stronger than the second. Diphthongs may therefore be described as centring (i.e. gliding towards the central vowel /ə/ as in /iə , eə , ʊə / and closing (i.e. gliding from a relatively more open to a relatively more close vowel). Closing diphthongs fall into two groups: (i) those ending in /i/ as in /ei, ai, ɔi/, and (ii) those ending in /ʊ/ as in /əʊ, ɑʊ/.

Finally, vowels differ in terms of tenseness (i.e. the degree to which the muscles of the lips and tongue are tightened (tense) or relaxed (lax). Lax vowels include /i, ɛ , æ, ə , ʊ , ʌ ,/ and tense vowels /i:, ei, ɑ ,

ɔ , əʊ , u:, ai, aʊ, ɔi /. While vowels may be identified on the basis of the features discussed above, the most important seem be those associated with vowel quality: tongue height, tongue advancement and lip shape.

6.2.1.3 Syllable structure

Every English word has at least one syllable, with as many as four or even more occurring in a number of words. Each syllable consists of a nucleus or PEAK which may be preceded and followed by one or more non syllabic elements, i.e. consonants. Those preceding the nucleus are referred to as the ONSET and those following as the CODA.1 The nucleus consists mainly of a vowel though any of the following consonants /m, n, ŋ , l, r/ may become syllabic in certain situations. In words such as "kitten" [kitn̩] and "candle" [kændl̩] the syllabic [n̩] and [l̩] occur in final position.

English has a highly complex syllable structure which may be summarised as follows:

$$(C_{0-3}) \quad V \quad (C_{0-4})$$

According to this schema, a syllable may consist of a vowel alone, as in "I" /ai/; it may have one to three

---

1 In some models of non linear phonology the term CODA is not used in the description of syllable structure. Rather the syllable is described as having an ONSET, a RHYME and a NUCLEUS which branches from the RHYME. The RHYME may also have a non nucleic element following the NUCLEUS (usually a consonant). See Chapter Three, section 3.5.

consonants before it as in "pie" "tree" "stray" /p_, tr_, str_/, and up to four consonants after it as in "add", "bask", "next", "prompts" /_d, _sk, _kst, _mpts/. English therefore allows cluster sequences of up four consonants. Other more complex syllable combinations can be derived from the elaboration of the above canonical structure.

However, there are constraints as to the types of consonant sequences that can occur in English.2 For example, neither /ʒ/ nor /ŋ/ may initiate an initial CV_ structure. /s/ does not occur before /r/ in a CC_ cluster, neither do /t, d, θ, ʃ/ before /l/ nor /p, b, f, ʃ/ before /w/. Initial clusters of three consonants are of the following types: sCr_, sCl_ and sCw_ where "C" represents any fortis stop.

In the final position consonants not allowed in the _VC structure include /r, h, j, w/. The following two consonant cluster types may occur: _rC, _lC, nasal or fricative + stop, stop + fricative and stop + stop. Three cluster sequences include nasal + two obstruents, liquid + stop + fricative and two liquids + stop. Final CCCC clusters occur as a result of the suffixation of a /t/ or /s/ morpheme. Up to four consonants are permitted in intervocalic position. These include clusters found in the initial and final positions as well as cluster combinations not occurring in these positions.

---

2 Refer to the notion of sonority sequencing discussed in section 3.6.2 of Chapter Three.

6.2.1.4 Suprasegmentals

The following suprasegmental features are considered: stress , intonation and tone. **Stress** applies to whole syllables in words occurring in isolation or in connected speech. Syllables may be stressed or unstressed. The former involve the use of greater muscular energy, and are more prominent than their unstressed counterparts, prominence being measured by the presence of a high pitch, increased length and loudness.

In English stress forms an essential component of a word's phonological form, since words become easily unrecognizable if the stress is wrongly placed. It therefore serves a variety of functions (Ladefoged, 1982). Firstly, stress serves to emphasis a word or contrast it with another. Secondly, it is used to indicate the syntactic relationships between words or parts of words.The stress on the first syllable of "'import" ['impɔ:t] distinguishes it from its verb form "im'port" [im'pɔ:t] which has stress on the second syllable. Thirdly, stress also helps to distinguish between a compound noun, e.g. 'hot dog (food) and an adjectival phrase 'hot 'dog (an overheated animal) (Ladefoged, 1982:105).

While the stress patterns of English remain systematic, they vary from word to word and from sentence to sentence. Consequently, it is difficult to predict on

the basis of the number of syllables in a word, which syllable the stress will fall on (O' Connor, 1973).

Intonation may be defined as the use of pitch over whole utterances - i.e whole phrases and sentences as opposed to words - to convey linguistic information (Ladefoged, 1982; O'Connor, 1973). Like stress, intonation serves a range of functions namely attitudinal, accentual, grammatical and discourse (Roach, 1983: 136/137). Through intonation we are able to express our emotions and attitudes, as well as contrast and emphasis. In the syntactic domain intonation helps to convey information regarding the placement of boundaries between phrases, clauses or sentences, and the distinction between declarative and interrogative sentences. The former are said to have a rising intonation and the latter a falling intonation. Lastly, in the discourse domain intonation may serve to focus the listener's attention on aspects of the message that are important, or to regulate conversational behaviour (Roach, 1983:147).

When pitch variations are used to distinguish the meanings of words, they are called **tones**. English is not a tone language in the sense described here. Anaang, on the contrary, constitutes a typical tone language to which we now turn our attention.

6.2.2. Anaang

In Chapter Five attention was drawn to the paucity of publications on the Anaang language (see for example Essien, 1970, 1970/71). In order therefore to provide a reliable description of its sound system we are forced to rely on additional sources drawn from available descriptions of related languages, i.e. Efik and Ibibio. Such sources include Akinlabi and Urua (1990), Connell (1989, 1991), Cook (1969/82, 1985), Essien (1983, 1986, 1990), Urua (1990).

While it may not seem like good linguistic practice to extrapolate from other languages in order to describe another, it is worth re-emphasizing the fact that these are highly mutually intelligible languages which share a lot of structural similarities (cf. Connell, 1991). However, we do not imply an absence of differences between them. Such differences will be highlighted as deemed necessary. Moreso, the present writer being a native speaker of Anaang will bring intuitive knowledge to bear upon the description.

6.2.2.1 Consonants

There are fifteen to nineteen phonemic consonants in the phonological inventory of Anaang depending on the dialect considered. For the Ukanafun dialect we have opted for nineteen as shown in the chart below. The liquids /r/ and /l/ have a special status and have therefore been parenthesized. From the point of view of

distribution they act as allophones of the same phoneme. However, the difficulty lies in assigning them to a particular phoneme. We shall return to this problem later. For the time being it is worth noting that in comparison with English, Anaang has a smaller phonemic inventory.

Seven oral **stop** phonemes are distinguished in five places of articulation. Unlike English, Anaang has labialized /g^w/ as well as labio-velar /k͡p/ plosives which constitute single phonemes rather than sequences of phonemes. The velar /k/ and labio-velar stops have no voiced counterparts, while the labialized stop exists in its voiced form alone. Interestingly, /g^w/ occurs in most environments which take the labio-velar approximant /w/ in Ibibio and Efik, i.e. word initially and after a vowel prefix. Examples include "wèt" [wèt], "íwát" [íwát] (Ibibio, Efik), "gwèt" [g^wèt], "ígwát" [íg^{w}át] (Anaang), meaning "write" and "grey hair" respectively. Anaang stops are examined in greater detail in section 6.3.4.

The **fricative** inventory is greatly reduced consisting only of the voiceless fortis labio-dental and alveolar pair /f, s/. They occur word initially, and after a vowel or nasal prefix as in the following examples: "fíp" [fíp-] "suck out", "sìp" [sìp-] "small", "àfò" [àfò] "you", "ùsèm" [ùsèm] "language", "ḿfòt" [ḿfòt-] "peel" (noun), "ńsuñ" [ńsuŋ] "fly" (noun).

Figure 6.3: Chart of Anaang consonant phonemes

Place of articulation

Manner of articulation	B/bl	L-dtl	Alv	Pal-alv	Pal	Vel	Lab'd vel	L/bl vel	Uvul
Plos	p b		t d			k	g^w	kp	
Fric		f	s						
Affric				tʃ	dʒ				
Nas (plos)	m		n		ɲ	ŋ			
Tap			ɾ						
C/tral Approx					j			w	ɣ
L/tral Approx			l						

N.B: B/bl = bilabial; L-dtl = labio-dental; Int-dtl = inter-dental; Alv = alveolar; Pal-alv = palato-alveolar; Pal = palatal; Vel = velar; Lab'd vel = labialized velar; L/bl vel = labial velar; Uvul = uvular.

The phoneme /ɣ/ while being classified as an uvular approximant is usually treated along with fricatives because it is often auditorily close to the velar fricative /x/ or /ɣ/ (Essien, 1990). Its two allophones include (a) [ɣ] which occurs between two back vowels as in "fághá" [fáɣá] "become tight", and (b) [R] a uvular tap or flap which appears to occur between front vowels as in "bèghé" [bèRé] "arrive" (e.g. at a port). [ɣ] and [R] contrast with [kk] an allophone of /k/

as follows: "fághá" / "fákká" [fákká] "uncork", "bèghé" / "bèkké" [bèkké] "belch".

Like English, Anaang has voiced and voiceless **affricates**, though in Anaang the voiced counterpart is more commonly a palatal /dʒ/ rather than a palato-alveolar /dʒ/ affricate (Connell, personal communication). However, there may be variation within and across speakers with regard to the use of /dʒ/ and /dʒ/. Affricates occur initially, and after a syllabic vowel or nasal. E.g. "chiọ́ñọ́" [tʃiɔ́ŋɔ́] "tighten", "jiọ̀ñọ́" [dʒiɔ̀ŋɔ́] "warm slightly", "íchíp" [ítʃíp-] "palm kernel", "újòm" [údʒòm] "noise" "ǹchát" [ǹtʃát-] "dried", "ńjèm" [ńdʒèm] "I want".

Both /dʒ/ and the voiceless /tʃ/ have /j/ and /s/, respectively, as counterparts in Ibibio, Efik and the Ikot-Ekpene dialect of Anaang. However, in Ukanafun Anaang /tʃ/ may alternate with /s/ in certain lexical items, in which case they occur in free variation. For example, "see" may be rendered as "sé" [sé] or "ché" [tʃé]. The occurrence of /s/ in such cases is often attributed to influence from speakers of related languages. In other cases /tʃ/ contrasts with /s/ as in "chin" [tʃìn] "refuse" (verb) and "sín" [sín] "put". The differential patterning of these segments may be traced to the process of sound change (see Connell, 1991).

Nasals in Anaang are the same as in English except that the velar nasal /ŋ/ has two allophones [ŋ] and [ŋ^{w}] a labialized counterpart, and the alveolar /n/

serves a syllabic function initially rather than finally. /ŋ/ is realized as [ŋ^{w}] initially and after a syllabic prefix as in "ñwàn" [ŋ^{w}àn] "dry" (verb), "ìñwáñ" [ìŋ ^{w}áŋ] "farm". It becomes [ŋ] word finally, or intervocalically provided the preceding vowel is not a prefix (Essien, 1990), e.g. "âbáñ" [ábáŋ] "pot", "dọ́ñọ́" [dɔ́ŋɔ́] "be ill".

While the syllabic nasal is phonemically /n/ it takes different orthographic forms being written as "m", "n" or "ñ" depending on the nature of the following consonant. These forms reflect an on-going assimilation process at the phonetic level, i.e /n/ assimilates to the following consonant becoming homorganic with it. For example, /n/ is realized as [n] before alveolar consonants [n], [d], [s] and [t], as [m] before labial consonants [b], [k͡p], as [ŋ] when followed by velar [k] or labialized velar [g^{w}] segments, and as [ɱ] before the labio-dental fricative [f]. Its status before affricates is uncertain.

Approximants in Anaang include the **glides** /j/ and /w/. The occurrence of /j/ is limited and its distribution restricted to the intervocalic environment in words such as "béyé" [béjé] "pass", "kpáyá" [kpájá] "be sour". Comparatively, the labio-velar approximant /w/ has a wider distribution in that it occurs initially, as well as after a syllabic prefix, e.g. "wíp" [wíp-] "ask", "Àwàsì" [àwàsì] "God". /w/ has the bilabial stop /b/ as its counterpart in Ibibio, Efik, and Ikot-Ekpene dialect of Anaang. For example, "àwǔ" [àwʊ̀] "prawns" is realized

as "ọ̀bŭ" [ɔ̀bŭ] (Ibibio, Ikot-Ekpene Anaang) and "ò bŭ" [òbŭ] (Efik). However, these differences are restricted to certain lexical items, as other "w" words such as "úwém" (life), "úwâ" (sacrifice) have no "b" equivalents, and remain unchanged across the languages. The liquid consonants consist of the alveolar lateral approximant /l/ and the alveolar tap /r/ which are examined in section 6.3.2.

6.2.2.2 Vowels

Anaang has a twelve vowel system made up of seven pure vowels /a, e, i, ɔ , o, u, ʉ/ and five diphthongs /ai, ei, ɨ, oi, ui/ (see chart below). The pure vowels have geminated or double counterparts which are phonemic in status, and serve lexical as well as grammatical functions (Essien, 1970). Geminate vowels contrast with the pure vowels as the following examples show:

bàk	"be early"	kóóñ	"hang"
bààk	"be afraid of"	wʉ̀k	"inform"
dép	"buy"	wʉ̀ʉ̀k	"bury"
déép	"scratch"	fut	"be noisy"
Kóñ	"cough"	fúút	"swell"

From the grammatical perspective double vowels serve to indicate plurality as in

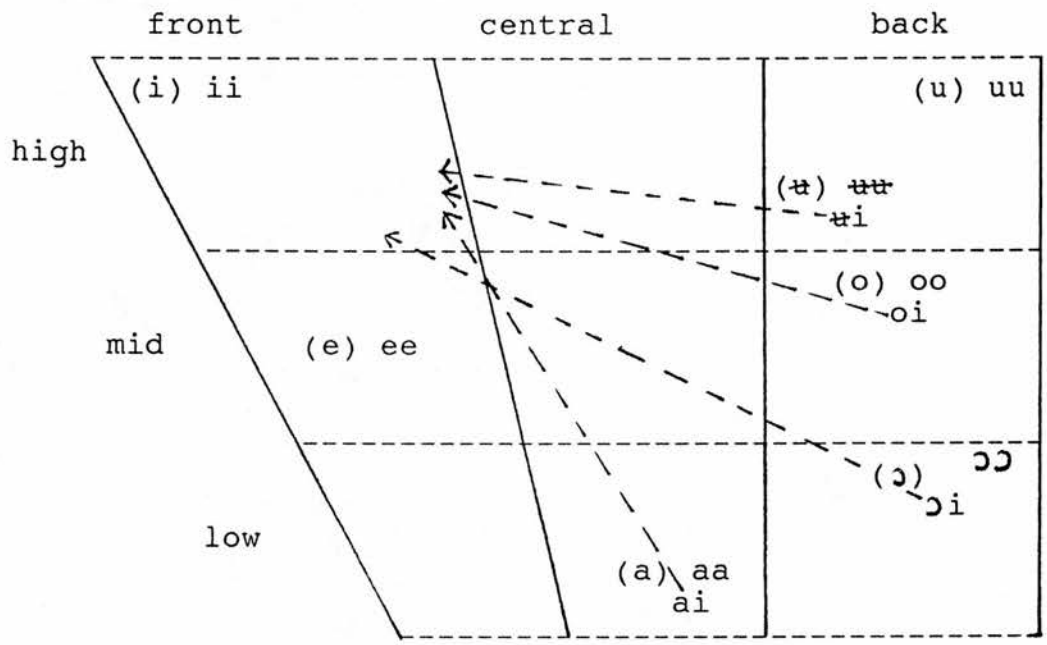
dip "hide something"

diip "hide many things"

bòm "break something"

bòòm "break many things"

Figure 6.4 Chart of Anaang vowels.



Short vowels are shown in parentheses. Dotted arrows indicate the gliding position of diphthongs.

It is worth noting that in spite of their phonemic status double vowels are often treated along with pure vowels. What is more, the process of vowel lengthening applies to Anaang and Ibibio, and not to Efik. Thus, in Efik "bom" (see above) expresses singularity as well as plurality, and "bak" has the two meanings indicated above.

The pure vowels consist of two front and one central unrounded vowels, as well as four back rounded vowels. The two high back vowels /u/ and /ʉ/ are phonetically similar, except that /ʉ/ is more open and has some degree of centralization. Of the seven pure vowels, four have no counterparts in English. They include (i) /a/ a low central vowel which is often substituted for English /æ/ in interlanguage performance, (ii) /ɔ/ which though has the same symbol as the English /ɔː/ is phonetically different in that it is a low back vowel and has shorter duration, (iii) /o/ which resembles French /o/ as in "beau" [bo], and (iv) /u/. Anaang, on its part has no counterpart for the English central vowel /ə/.

Apart from /ɔ/, /o/ and /ʉ/ which occur only medially and finally, all the other pure vowels occur in every position. Geminated vowels are found only medially in closed syllables (CVC), while diphthongs occur only in open syllables (CV). Diphthongs may be described as closing since they involve a glide towards /i/. They occur in the following words: "dái" [dái] "lick", àjèi [àdʒèi] "young furled shoot of palm tree", "kói" [kói] "scoop", "kwọ́i" [kwɔ́i] "peel" and "gwʉ́i" [gʷʉ́i] "peg" (verb).

6.2.2.3 Syllable structure

Syllables in Anaang have the basic canonic structure CV(C) which may be elaborated upon by the

addition of a syllabic nasal (N-) or vowel (V-) occurring as prefixes mainly in nouns. See the following examples:

CV		V-CV N-CV	
dí	"come"	á lú	"hole"
kpà	"cut"	ń sá	"game"

CVC		V-CVC N-CVC	
tóp	"throw"	íják	"fish"
bàt	"count"	m̀fòt	"peel"

The canonic structure may also be elaborated upon by the addition of other combinations of C and V. Thus, the following CVCV (béñé "beg"), CVCCV (dèppé "emerge suddenly") or CVVC (fíík "press") are possible syllable structures. Words consisting of four syllables or more also occur in the language. It is worth noting that while Anaang allows intramorphemic consonant geminates as in the CVCCV structure cited above, it disallows the occurrence of dissimilar CC sequences in the same syllable. In other words there are no consonant clusters in Anaang in the sense that there are in English. Anaang therefore constitutes a less marked system in this regard (see Greenberg, 1966). Any consonant phoneme may occur as the ONSET of the syllable, while the restricted set of plosives and nasals /p, t, k, m, n, ŋ/ may occupy the CODA.

6.2.2.4 Suprasegmentals

While **stress** plays a major role in English as indicated in section 6.2.1.4, its functions in Anaang (as in Ibibio and Efik) are limited to the expression of emphasis, contrast and emotion (irritation, anger, frustration or desperation) (Essien, 1990). In the same vein, **intonation** plays a minor role in the language serving only certain syntactic/semantic functions such as the expression of modality - the speaker's attitude or involvement in what he/she says.

Tone, on the contrary, constitutes the most important suprasegmental feature in Anaang phonology. The basic tones are "high" and "low" represented by the symbols (◌́) and (◌̀) respectively. They may combine on a given syllable to give "falling" (◌̂) and "rising" (◌̌) tones, both of which are contour tones (i.e. tones involving gliding movements). Throughout this chapter we have assigned tones to all Anaang words used as examples to illustrate certain phenomena. However, to distinguish clearly between the tones further examples are given below:

dèp	"rain" (verb)	low
dép	"buy"	high
úwâ	"sacrifice"	falling
ìmǒ	"wealth"	rising

A phonologically significant tone referred to as "downstep" (ꜜ) results from the slight lowering of a high tone following another high tone as in "údèꜜ" "there", "ḿbòꜜk" "please". This tone differs from and contrasts with the basic tones.

Tones function both lexically and grammatically. From the point of view of lexis they serve to distinguish meanings of words. For example "ákpán" "first son" differs from "ákpàn" "basket", and "únèn" "chicken" from "únén" "justice". Like stress in English, tones form an essential component of the phonological form of Anaang words. As such, words become easily unrecognizable if the wrong tones are used.

From the grammatical perspective tones serve to indicate person and number. In Anaang verbs usually have concord markers which indicate the person or persons carrying out an action. Consequently, these verbs are not necessarily preceded by corresponding personal pronouns in sentences. The first person concord marker which is always a nasal /n/ carries a high tone. Other persons have vocalic concord markers. Where the same vowel represents more than one person the tones differ. For example, the second person singular marker is distinguished from the third person by the presence of a low tone. The third person, on the contrary, always carries the high tone. Consider these examples:

Àdià ǹbòró "You are eating a banana."

Ádià ǹbòró "He/she is eating a banana."

The phenomenon of tones functions in more complicated ways, the discussion of which goes beyond the scope of the present study. However, for a more detailed examination of tones see Cook (1985).3

6.3 Selected phonological segments

Having considered the sound systems of English and Anaang from a general perspective we turn our attention to the specific structures to be investigated. A detailed description of liquid and stop segments is presented below.

6.3.1 Liquid segments in English and Anaang

As was indicated in section 6.2.1.1 English phonology distinguishes two **liquid** approximant phonemes transcribed as /l/ and /r/, and marked by the presence or absence of the feature *lateral*. /l/ is realized with the tongue tip on the alveolar ridge, a process which blocks air passage in the middle of the oral cavity, allowing it to escape along one or both sides of the tongue, i.e. unilaterally or bilaterally.

In English, three phonotactically conditioned allophonic variants of the alveolar lateral /l/ exist -

3 Though Cook focuses on Efik many of the phenomena described are applicable to Anaang.

"clear", "dark" and "voiceless approximant" (Gimson, 1980). The clear [l] precedes vowels word initially and medially as in "leave" and "silly", and forms clusters with voiced stops /b, g/ in the said positions as in "blow" and "glow". The dark [ɫ] occurs word finally after vowels as in "feel" and "call", and before consonants in final cluster position as in "help" and "salt". Its role as a syllabic sound following a consonant has already been discussed (see section 6.2.1.3). The above distributional patterns indicate that both variants of /l/ exist in complementary distribution. The voiceless variant [l̥] occurs in clusters after /p, t, k, s, f/ as in "play", "butler", "clean", "wrestler" and "flow".

The non lateral liquid /r/ is realized phonetically as a voiced alveolar approximant [ɹ], the airstream escaping over the centre of the tongue. It has a tapped variant [ɾ] which occurs intervocalically in words like "very" and "sorry", where /r/ is preceded by a short stressed vowel and followed by an unstressed one (O'Connor, 1973). This variant also occurs after voiced and voiceless interdental fricatives as in "three" and "withering". However, the alveolar tap is not a common phenomenon in RP. It occurs more predominantly in non RP forms of English (e.g. some varieties of Scottish English), and most forms of American English (Ladefoged, 1982). In the case of American English, the tap occurs as an allophone of /t/ or /d/ intervocalically as in "better" or "daddy".

RP /r/ has a less complex distributional pattern than /l/ in that it precedes but does not follow vowels, except under certain conditions, as we shall see below. It occurs word initially and medially as in "read" and "dairy", and forms clusters with voiced as well as voiceless obstruents except /s, z, tʃ, dʒ/ in the two positions. Like /l/ it becomes voiceless following /p, t, k, f, ʃ/. It rarely occurs in final position except as a linking "r" before a word beginning with a vowel, e.g. "far away", "for all" , or as an intrusive "r" where it is inserted before a following vowel even though there is no "r" orthographically as in "China and" (Trudgill and Hannah, 1982: 14).

Unlike in English where /r/ and /l/ are easily classified into neat categories, establishing the phonemic status of liquid segments in Anaang is not so straightforward given their patterning in the language. Phonetically, the lateral segment is an alveolar approximant [l], and the non lateral an alveolar tap [ɾ]. Sometimes a trill occurs in emphatic or very careful speech. Like most African languages with a lateral liquid, Anaang has the clear variety which occurs word initially and after a syllabic prefix (i.e. stem initially) as the following examples indicate:

lómó	[lómó]	"measure"
lád	[lád]	"take"
álàn	[álàn]	"oil"
élìm	[élìm]	"rain" (noun)

The lateral does not occur finally in Anaang.

The alveolar tap, on the contrary, is restricted to the intervocalic (non syllabic prefix) position as in

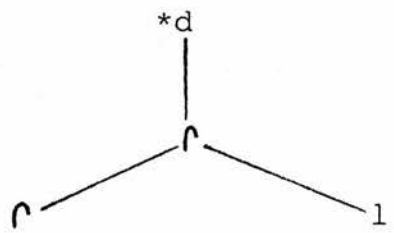
m̀bàkárá	[m̀bàkáɾá]	"a European"
bóró	[bóɾó]	"be round"
kúfìré	[kúfìɾé]	"do not forget"

Given the above distributional patterns it is obvious that the alveolar lateral and tap occur in complementary distribution. This applies to the Ukanafun (UK) dialect of Anaang which distinguishes [ɾ] and [l] phonetically but not phonemically. Other dialects lack the alveolar lateral, but have the tap either as a phoneme /ɾ/ (e.g Abak Anaang - AA) or as a weakening of /d/ (e.g. Ikot Ekpene Anaang - Ik). In certain variations of UK Anaang, and in the speech of certain individuals [ɾ] and [l] may occur in free variation in initial and post-syllabic prefix positions, and may occasionally alternate with /d/. In rare cases the intervocalic position may exhibit [ɾ]-[l] alternations.

The question then arises as to which phoneme to assign these variants. Recall that [ɾ] and [l] occur in

complementary distribution in the main UK dialect. While the choice of a phonemic symbol remains arbitrary, we propose /l/ as the phoneme given its wider phonological distribution and the fact that native speakers of Anaang usually associate this segment with the Ukanafun dialect. Moreso, there is some evidence which seems to assign an important role to /l/ in the dialect.

Relying on evidence from electropalatography, Connell (1991) suggests that the occurrence of /l/ in UK Anaang may have resulted from gradual sound change which affected alveolar segments /d, t, l/ in languages of the Lower Cross group (see Chapter Five). With regard to Anaang it would seem that the change followed the pattern *d > [ɾ] > [l] as illustrated below:



According to this schema the proto-alveolar *d became [ɾ] through weakening, and the latter underwent a further split into [l] in the UK dialect. These changes seem to have affected the initial and post-syllabic prefix positions. It will be recalled that in the examples cited above [l] occurs in these two positions. However, it must be borne in mind that these observations are

based on a highly experimental study. A consideration of sociolinguistic factors may well reveal other trends.

Given the above description of liquid segments the main difference between English and Anaang, therefore, lies in the maintenance of phonemic contrast in the former, and the non maintenance of such contrast in the latter. We have also noted phonetic differences regarding the realization of the non lateral liquid in the two languages: English has an approximant [ɹ] and Anaang a tap [ɾ].

According to Maddieson's (1984) classification of liquid segments based on a survey of 317 languages, approximants constitute the most frequent (80%) lateral liquid, while taps (38%) come second to trills (48%) among the non laterals. The approximant [ɹ] is the rarest constituting 14% of all non lateral liquids. A markedness relationship based on frequency may therefore be established between the latter group such that approximants are more marked than taps which are in turn more marked than trills. By implication English constitutes a more marked system than Anaang. What remains to be seen is whether this markedness relationship has any bearing on the L2 acquisition of /r/. With regard to lateral segments the approximant is the least marked.

The description of /r/ and /l/ and their distribution in the two languages may be summed up in the following realization rules:

English

/r/ ---> [ɹ]/ #__; #C__; V__V; __C#; __#

/l/ ---> [l]/ #__; #C__; V__V

[ɫ]/ __C#; __#

Anaang

/l/ ---> [ɾ] #__; #V__V; __V#; V__V

[l]/ #__; #V__V

where "#" indicates 'word boundary' (initial or final), "C" means 'consonant', "V" vowel and "/" 'in the environment of'. No realization rules have been written for intervocalic alveolar tap in English since we consider /r/ to be predominantly an approximant phonetically.

6.3.2 Oral stop segments in English and Anaang

Bilabial /p, b/, alveolar /t, d/ and velar /k, g/ have also been selected for consideration in the present study. The three stages of plosive articulation include: (i) the closure stage during which the two articulatory organs come together to form an obstruction in the oral cavity, (ii) the hold stage during which the compressed air is stopped from escaping, (iii) the release stage during which the articulators part to release blocked airstream (Abercrombie, 1967; Gimson, 1980; Roach, 1983).

The airstream is released with or without a small burst of sound (i.e. aspirated or unaspirated).

In Anaang (as in Efik and Ibibio) the third stage of plosive articulation does not take place when stops occur word finally (Connell, 1991; Essien, 1990; Urua, 1990). The articulatory gesture for word final stops therefore remains incomplete, and the stops are unreleased. Final unreleased stops often give the illusion of being "voiced", but phonetically, they are clearly voiceless. We can therefore argue for the neutralization of voicing contrast in Anaang stops word finally in favour of the voiceless segment. However, if final stops are immediately followed by a vowel, then they are not only released, but weakened to voiced homorganic fricatives, taps/trills and approximants as will be demonstrated below.

Stops may be released or unreleased word finally in English. Released and non released stops therefore occur in free variation in the said position. The length of the preceding vowels usually distinguishes voiced from voiceless stops (Gimson, 1980), shorter vowels occurring before voiceless /p, t, k/ and longer vowels before voiced /b, d, g/. When two stops occur in succession, either within words or at word boundaries (e.g. apt, up to), the first stop is not released until the second is made. Gimson (1980) points out that the non release of final stops is a feature of colloquial RP, since careful speakers tend to release final stops , especially in

formal conversations. Thus, while the non release of final stops functions obligatorily in Anaang, this does not seem to be the case in English.

Of the six plosives identified above, only five occur in the phonemic inventory of Anaang: the voiced velar stop /g/ does not exist in the language. A survey of the world's languages reveals that among languages which exhibit a voicing contrast in the stop series, but have one stop missing, the missing element is typically a /g/ (or a /p/) (Locke, 1983). The absence of the voiced velar stop /g/ therefore results in a half symmetric system (Essien, 1990). The phonemes are displayed below in the two languages:

English	Anaang
p t k	p t k
b d g	b d -

We shall now consider the stops individually, describing their phonetic characteristics as well as their distributional patterns.

6.3.2.1 Bilabial stops: /p, b/

Bilabial stops are articulated with the upper and lower lips, and they exhibit the fortis/lenis, voicing and aspiration characteristics discussed in section 6.2.1. In English /p/ and /b/ contrasts in all positions either as singleton consonants or as cluster sequences

with other consonants. Only the voiceless fortis stop follows /s/ in an initial cluster sequence occurring in the same syllable, in which case it does not undergo aspiration. In initial prevocalic position, however, /p/ is fully aspirated. Full voicing for /b/ can only be achieved intervocalically. The segment remains partially voiced in initial and final positions.

In Anaang bilabial stops do not contrast word initially (except in the Ikot Ekpene dialect). The Ukanafun dialect displays no /p/ in initial position; instead /p/ occurs intervocalically, realized as a geminate [pp] and finally, realized as [p-], a voiceless unreleased allophone when not followed by a vowel. However, at word boundary before a vowel /p/ undergoes weakening to a voiced bilabial fricative [β]. The three allophones of /p/ are illustrated in the following examples:

kọ̀ppọ́	[kɔ̀pɔ́]	"disentangle"
tóp	[tóp-]	"throw"
tóp # étó	[tóβétó]	"throw a stick"

The phoneme /b/ occurs initially, after a syllabic prefix and medially where it is weakened to voiced bilabial fricative [β], or occurs as a geminate [bb] as in the following examples:

bàt	[bàt-]	"count"
íbàt	[íbàt-]	"mathematics"
dóbbó	[dóbbó]	"be heavy"
díbé	[díβé]	"hide"

Unlike its English counterpart, Anaang /b/ undergoes full voicing word initially. That is, the vocal cords continue to vibrate throughout the entire duration of /b/, and not just during the release phase as is the case in English. The geminate [bb] alternates freely with [pp] as in "dóbbó" and "dóppó" which both mean "heavy". In other words, [pp] and [bb] are not phonologically contrastive. The choice between the two geminate allophones varies across speakers, and to an extent across dialects. [bb] seems to occur more frequently in the UK dialect of Anaang, while [pp] is the preferred variant in the IK dialect. However, individual variation probably constitutes a stronger factor than dialectal variation.

When [β] occurs as an allophone of /p/ and /b/ at word and morpheme boundaries respectively, the problem of syllabification arises. The solution proposed by Cook (1969/82) is to assign [β] to the preceding as well as following syllables, in which case it becomes an ambisyllabic consonant. As the above examples indicate, the latter contrasts with geminate consonants. The following rules reflect the distribution of /p/ and /b/ in English and Anaang.

English

/p/ ---> [p^h]/ #__; V__V

[p]/ __#

/b/ ---> [b]/ #__; V__V; __#

Anaang

/p/ ---> [pp]/ V__V

[p-]/ __#

[β]/ V__#V

/b/ ---> [b]/ #__; #V__V

[bb]/ V__V

[β]/ V__V

where [p^h] means "aspirated /p/". It will be noted that only distributional rules considered to be essential have been written for English.

6.3.2.2 Alveolar stops: /t, d/

Alveolar stops are realized with the tongue tip or blade touching the alveolar ridge. In English /t/ and /d/ occur in all contexts as singleton consonants and in cluster sequences with other consonants. The voicing characteristics for /d/ are the same as for the bilabial plosive /b/. Like its bilabial counterpart, /t/ is fully aspirated initially. No such aspiration occurs following an /s/ in a cluster sequence. The fortis alveolar stop

has an additional variant - the glottal stop [ʔ] - which occurs restrictively in RP.

The articulatory characteristics of alveolar stops differ slightly in Anaang. Given the wider area of contact between the tongue blade and the alveolar ridge, /t/ is better described as laminal denti-alveolar (or post-alveolar) (Connell, 1991). /d/ on the contrary has a smaller area of contact and may be more specifically described as lamino-alveolar. A slight affrication of /t/ is heard before the high front vowel [i]. The aspiration of Anaang /t/ word initially is relatively weak in comparison with English /t/. Like its bilabial counterpart, /d/ undergoes full voicing in initial position.

Alveolar segments contrast phonemically in the two languages, but the contrast appears to be restricted to word initial position in Anaang. Geminate allophones of /t/ --> [tt] and /d/ --> [dd] alternate intervocalically mainly in verbs. Again individual or dialectal preferences determine the choice between the two allophones. [dd] would seem to be the preferred variant in UK Anaang. In final prepausal position /t/ remains unreleased. However, at word boundary before a vowel /t/ weakens to [ɾ] a voiced alveolar tap. The tap may also occur as a weakening of /d/. E.g.

támmá	[támmá]	"jump"
dámmá	[dámmá]	"become mad"

ìtiěn [ìtjěn] "five"

ídáp [ídáp-] "sleep"

sítté [sítté] "uncork"

síddé [síddé] "uncork]

ḿkọ́t # Èkpènḗ [íkɔ́ɾèkpènḗ] "place name"

The following rules apply to the distribution of /t/ and /d/ in English and Anaang:

English

/t/ ---> [tʰ]/ #__; V__V

[t]/ __#

/d/ ---> [d]/ #__; V__V; __#

Anaang

/t/ ---> [t]/ #__; V__V

[tt]/ V__V

[t-]/ __#

[ɾ]/ V__#V

/d/ ---> [d]/ #__; #V__V

[dd]/ V__V

[ɾ] / V__#V

6.3.2.3 Velar stops: /k, g/

Velar stops are realized when the back of the tongue touches the soft palate (velum) which has been raised to block the nasal resonance. Both stops exhibit the same distributional patterns as the other stops in

English. /k/ remains aspirated in initial position, while /g/ undergoes partial voicing initially and finally. They also contrast phonemically in English.

As already indicated the phoneme /k/ has no voiced counterpart in Anaang, though /g/ may be found in loanwords such as "Ugep", "Lagos" (both Nigerian cities). The voiceless velar stop occurs initially, after a syllabic prefix, medially as a geminated [kk] in verbs and nouns derived from verbs, and finally as an unreleased stop (Essien, 1990). At morpheme boundary, before a vowel /k/ is weakened to a voiced uvular approximant [ɣ], tap [R] or velar fricative [ɣ].

ká́rá́	[kaɾa]	"govern"
ú́kó́t	[ú́kó́t-]	"leg"
dàkká́	[dàkká́]	"move away"
ké́k	[ké́k-]	"fell"
ké́k # étó́	[ké́Ré́tó́]	"fell a tree"

The following rules summarise to the distribution of /k/ and /g/ (where necessary) in the two languages.

<u>English</u>

/k/ ---> [kʰ]/ #__; V__V

[k]/ __#

/g/ ---> [g]/ #__; V__V; __#

Anaang

/k/ ---> [k]/ #__; #V__V

[kk]/ V__V

[k-]/ __#

[R]/ V__#V

[ɓ]

6.4 Summary and conclusion

We began this chapter by identifying the models of English and Anaang which would serve as the norm for the purpose of our study. It was argued that while accessibility to the RP accent remains open to a small number of English speakers, the model serves as the **ideal** (even if in theory) in most language teaching and learning circles. Hence, it constitutes the norm in most contrastive studies. An overview of the sound systems of English and Anaang was then presented, followed by a comparison of liquid and stop speech sounds in the two languages. Crosslingual differences and similarities in the phonetic/phonemic organization of the said speech sounds were highlighted. What implications these differences and similarities have on acquisition will be seen in the analyses of the results in Chapters Eight and Nine, and in the discussion of these results in Chapter Ten. Meanwhile, the next chapter focuses on the design of the study.

CHAPTER SEVEN: EXPERIMENTAL DESIGN

7.1 Introduction

The issues discussed in the preceding chapters constitute the background against which the rest of the study is set. In this chapter we present a description of the research design, and the procedures followed in the data collection. We begin with the discussion of the methods of investigation. Details are given regarding the preparation of the data collection instruments, and the administration of these instruments during the field study, which includes a pilot phase. The study adopts a cross-sectional approach which entails the observation of the performance of learners representing a range of language proficiencies.

7.2 Methods of investigation

This section presents a description of the data collection instruments and how they came into being. It describes the manner of administration, and the subjects to whom they were administered during the fieldstudy.

7.2.1 Preparation of data collection instruments

All instruments for data collection were prepared in Edinburgh before the writer travelled to Nigeria to carry out fieldwork. For the purpose of this study two separate experiments were designed for liquid and stop

segments respectively. A set of test materials was constructed for each experiment. Initially each set consisted of two word lists (one English and the other Anaang), a sentence list, a short text and a list of topics for discussion. These were to serve for reading, speaking, listening and translation tasks. In the final administration of the tests following piloting the translation task was eliminated. Also eliminated from the study was the sentence reading task for the stop data (see section 7.2.2.1). There are therefore no descriptions of the said tasks in this section.

The tasks were intended to elicit data representing different speech styles ranging from formal to informal. The necessity of collecting data from as many styles as possible has been stressed by Tarone (1983), which is that it allows for the accurate identification of patterns within and across styles, as well as the relative influence of phonological processes.

7.2.1.1 Experiment 1: liquids

(a) Word list

As already indicated the word list was designed to elicit the most formal (or careful) speech style. Sixteen pairs of test words were selected such that four pairs contrasted /r/ and /l/ in four word positions: initial (WI), initial consonant cluster (ICC), intervocalic or medial (WM) and final post vocalic (WF)

(see Appendix 1). Eleven pairs were drawn from pre-existing word lists used by Gillette (1980), Mochizuki (1981), Sheldon and Strange (1982) who examined the acquisition of liquid contrasts by Japanese and Korean learners of English. The remaining five pairs were taken from *Cobuild English Language Dictionary* (1987).

Within each subset of words, vowels varied systematically on both front-back and high-low dimensions. There were monophthongs as well as diphthongs. For the four pairs of WM items two had stress on the syllable preceding the liquid, and the other two on the syllable following the liquid. All other test words were monosyllabic. The clusters included voiced/voiceless stop + liquid, as well as voiceless labiodental fricative + liquid combinations. In addition to the 32 test words there were six pairs of filler (or dummy) words, two pairs each contrasting phonemes in initial, medial and final positions. Altogether there were 44 test items, each written boldly on 6" x 4" cards of which four packs were prepared.

The speech stimuli for the perception task consisted of the 32 /r/-/l/ words which were recorded in the Phonetics Laboratory of the University of Edinburgh. The stimuli were mastered onto a Sony PCM-F1 BETAMAX Digital Recorder, and simultaneously copied onto a Marantz CP430 Cassette Recorder using high quality Maxell XL1-S cassettes. Two recordings of four repetitions of the items were produced, one by a native speaker of

Anaang (Speaker no. 1), and the other by a native speaker of English (Speaker no. 2). Each recording consisted of different randomizations of the test items (Appendix 5). The words were separated by a pause of approximately three seconds to allow for listeners' judgement, and long pause of eight seconds inserted between groups of words. Two answer sheets - A and B - were prepared for each randomized version (Appendices 6 and 7).

(b) Sentence list

The 44 test items were embedded in sentences to create a less formal speech style (Appendix 3). Though care was taken to ensure the accessibility of these sentences to all categories of subjects, it was observed during piloting that some sentences were particularly difficult for low proficiency learners. A total of eleven sentences were rephrased and/or shortened to make for easier reading. All revisions were incorporated into the main study. Each sentence was written on a 6" x 4" cards of which four packs were prepared.

(c) Text

A short text of roughly 150 words was chosen from John Munonye's *The Only Son*, one of many African novels recommended for reading in many Nigerian schools (Appendix 4). The text was intended to provide an even less formal speech style, something intermediate in degree of monitoring between sentences and spontaneous

speech. It had altogether 42 liquid items: 23 /r/ and 19 /l/. However, we could not control for the phonological environments of the items. For example, /r/ words appeared in three positions (initial, cluster and final), while /l/ items appeared in all positions including word medial. The text provided a range of clusters which included initial, medial and final along the voiced-voiceless dimensions. Vowels also varied on both front-back and high-low dimensions.

7.2.1.2 Experiment 2: stops

(a) Word list

Twenty four test words were drawn from the *Cobuild English Language Dictionary* (1987), and grouped according to their places of articulation: bilabial, alveolar and velar (Appendix 2). Each group consisted of eight monosyllabic/disyllabic words, half of which had voiced, and the other half voiceless final stops. Care was taken to select simple lexical items, though one item "handbag" proved to be difficult to read during piloting, and was consequently shortened to "bag". Within each group vowels also varied on both front-back and high-low dimensions. Ten filler words were added to the list, thereby bringing the total number of items to 34. Each word was written on 6" x 4" cards of which there were four packs.

(b) Text

A text of about 150 words was selected from Peter Abraham's *Mine Boy*, another African novel recommended for reading in Nigerian schools (Appendix 4). It had 22 final stop words in utterance medial as well as utterance final positions. These included 2 voiceless bilabial (/p/), 10 voiceless alveolar (/t/), 9 voiced alveolar (/d/) and one voiceless velar (/k/) items. Vowels varied along the front-back and high-low dimensions.

7.2.1.3 Oral interviews: conversations

The main reasons for including interviews in this study was to elicit natural speech data. However, we would not claim that the data were completely unselfconscious vernacular speech, as this kind of speech is not readily obtained in face-to-face interviews. Like Beebe (1987:300) we would argue that the data were "natural" in the sense that they were "appropriate to the occasion", i.e. to the interview which is a natural speech event in society.

Several topics were proposed by the investigator which required the subjects to (a) narrate a story, an event, (b) discuss their future careers or hobbies, (c) describe a procedure (e.g for making palm oil), (d) express an opinion on socio-political events in Nigeria or elsewhere. The topics are not exhaustive but cover a wide variety of interests.

A comparison of reading and conversation tasks show that while the former have the advantage of rigorously controlling the phonological environments of the target sounds as these remain constant for all subjects, the latter do not (Major, 1986a). In conversations the investigator cannot exercise "absolute" control over the environments in which the target sounds are produced.

7.2.2 The Field Study

The field study was carried out in Nigeria between September, 1991 and January 1992, and in Edinburgh in April, 1992. It consisted of a pilot phase and the main study itself.

7.2.2.1 The pilot phase

The **pilot** phase of the study aimed at assessing the practical aspects of administering the instruments, as well as the reliability and validity of the entire data gathering process, with a view to modification prior to the main study. As a result, the subjects' performance were not evaluated for the purpose of statistical analysis.

The test materials were administered to six native speakers of Ukanafun Anaang representing four levels of proficiency intended for the main study: low (2), lower intermediate (2), upper intermediate (1) and advanced (1). They were volunteer participants pursuing secondary

and university education in Uyo. Data were collected in three sessions lasting 45 minutes each on different days. The procedures followed were the same as those described in section 7.2.2.3 below.

The data were then transcribed by the investigator and an assistant who was a lecturer of Linguistics in the University of Calabar, Nigeria. A point to comparison of the transcriptions showed 100% agreement on the liquid data, and 80% agreement on the stop data. It was on the basis of these transcriptions and the observations made during piloting that some of the test materials were modified or even eliminated. Some of these revisions were highlighted in section 7.2.1.

The changes made affected the number, nature and contents of the tasks. The translation task, for example, was eliminated from the main study because it made too many memory demands on the subjects who found it difficult to access many of the target words in spite of the cues that were given to aid recall. The number of tasks was reduced from six to five. The sentence reading task for the stop data was also eliminated because it did not yield the results expected.

Further changes were made regarding the duration of tasks and number of sessions during which the main data collection would take place. It was decided that the duration of the conversation would be cut by half allowing subjects five minutes on each topic. We felt that ample data could be collected within this time

period. We also decided to reduce the number of sessions from three to one given the changes made. There was also an overriding practical reason for doing so. We wanted to rule out any cases of subject mortality or attrition, whereby subjects dropped out before the sessions were completed. All these revisions were subsequently incorporated into the main study which is described below.

7.2.2.2 The main study

(a) Subjects

Eighty speakers of Ukanafun Anaang and 20 native speakers of English served as subjects for the main study. They included 73 males and 27 females who were pursuing either secondary or university education. The Anaang speakers constituted four experimental groups representing the four levels of proficiency indicated above (see section 7.2.2.1). Native speakers of English made up the fifth proficiency group which served as the control group to provide comparative data to check with the experimental groups.

Non-native proficiency was measured by educational attainment in terms of the years of exposure to formal English instruction. As such University students who have had a longer period of formal instruction in English constituted the upper intermediate and advanced groups. Low and lower intermediate learners consisted of

secondary school students who have had a shorter period of formal instruction. There is general consensus among SLA researchers that the length of formal contact with English has a positive effect on proficiency (see Ellis, 1985a).

However, what they do not agree on is what constitutes the best measure of proficiency (Larsen-Freeman and Long, 1991; Odlin, 1989), given alternative measures such as the rating of subjects by native English speakers (Suter, 1976), or the use of standard Proficiency Tests. Intuitively one would opt for the latter which seem to provide a quantitative measure of proficiency. We do recognize the problem with using education as a criterion for determining proficiency (especially pronunciation). That is, that there are speakers who perform below or above their expected level of competence (as defined by educational attainment) (see Bamgbose, 1982). Indeed, this was the case for some of our subjects, though such cases were relatively few (see Chapters Eight and Nine). As such, the earlier claim made regarding the positive relationship between formal instruction and proficiency is not invalidated.

It must be stressed, however, that neither of the alternatives mentioned above were available to the researcher prior to conducting the study. Moreso, even if the researcher had access to a standard Proficiency Test, the administration of such a test would have been

problematic given the practical reasons that were stated in the last section.

Prior to selection, the subjects filled out a language background questionnaire (LBQ) which provided among other things, several gross measures of linguistic experience: L1 background, other language knowledge, years of formal English instruction, type of primary school attended, average use of English and years (if any) of residence in an English speaking country abroad (Appendix 8). Native speakers were only required to provide information regarding their age, sex and year of study. Information obtained from the questionnaires was used in subject selection. Certain categories of learners were therefore eliminated from the study in a bid to control for extraneous variables. Those eliminated included (a) students who had attended international primary schools where English is the sole medium of instruction, and (b) speakers of languages other than Anaang and English. The profile of the subjects is presented in Appendix 9.

(i) Low learners

The low learners comprised 20 participants randomly selected from among first year students of the junior secondary section of Holy Family College, Abak. This is an all male school selected through a random sampling of schools in the Anaang speaking areas of Akwa

Ibom State of Nigeria. The subjects ranged from age 11 to 14 years, with an average age of 11.8 years. They had six years1 of formal English instruction in primary school. At the time of investigation they were beginning their seventh year of instruction.

None of them reported having resided abroad. All but two subjects reported using English "sometimes" for interaction beyond the school environment. The other two subjects used it "often". The investigator was informed that English served mainly for interaction with speakers of other languages. It is worth noting that opportunities for such interaction are often infrequent given that the learners live in linguistically homogeneous communities where Anaang is the main language of communication.

(ii) Lower intermediate learners

The 20 lower intermediate learners were also randomly selected from among students of the above mentioned school. These were first year senior secondary school students, aged between 14 and 17 years, with a mean age of 15.25 years. At the time of investigation they were in their tenth year of English instruction, having completed 6 years of primary and 3 years of junior secondary education. None of them reported having lived

1 It will be recalled that in the first three years of primary education English is taught as a subject. It becomes the medium of instruction from the fourth year onwards.

abroad. They all reported using English "sometimes" for interactive purposes outside the school environment, particularly for inter-ethnic communication. They, like their low level counterparts, live in linguistically homogeneous communities where the need for interaction in English is minimal.

(iii) Upper intermediate learners

The upper intermediate learners were first year undergraduate students pursuing various programmes in the Universities of Calabar and Uyo in Nigeria. They were recruited through notices on campus bulletin boards and personal contact. The 20 students who met the selection criteria outlined above were picked from among several volunteer participants. There were 8 males and 12 females, aged between 18 and 24 years with an average age of 21.4 years.

They had had 12 years of English instruction (6 years each of primary and secondary education). As first year undergraduates they were in their thirteenth year of instruction. Six subjects each reported using English "sometimes" and "very often", while 8 used it "often". Interestingly, English was not restricted to inter-ethnic communication, but was used even between speakers who shared the same L1 - Anaang. There are several explanations for this trend, one of which is the increasing number of opportunities for using English:

wider social contacts, travel, living and studying in fairly heterogeneous urban communities.

(iv) Advanced learners

There were 20 volunteer participants, also recruited from the Universities of Uyo and Calabar. They comprised 14 final year students, who were aged between 23 and 34 years with an average age of 27.5 years. Altogether there were 15 males and 5 females. They had had between 16 to 18 years of English instruction depending on whether they were undergraduate or postgraduate students. None of them reported having resided abroad, and like their upper intermediate counterparts they used English frequently beyond the classroom environment. More than half (13) of them used English "very often", while 4 used it "often" and 3 "sometimes". The 6 postgraduate students had worked for about 1 to 3 years before returning to university.

(v) Native speakers

The native English speaking subjects, 10 males and 10 females, were recruited from the University of Edinburgh. There were 14 undergraduate and 6 postgraduate students, all pursuing various programmes in the university. They were aged between 19 and 41 years, with a mean age of 25.2 years. There were 19 subjects from England and 1 from Scotland. Four subjects had had extensive experience with non native speakers of English

while teaching ESL and EFL in Africa and Asia. The native speakers, as already indicated above, served as a control group to provide comparative data to check with the experimental groups.

(b) Procedures

This section describes the procedures that were followed in data collection. Given similarities in the nature of tasks, we used the same procedures for both liquid and stop data.

(i) Task one: citation

Each subject was given a randomized pack of cards containing a card for each word on the two lists. The investigator ensured that every subject had a different randomization of words in a bid to control for the effect of order. Firstly, they read through the entire list silently in a familiarization session, during which they were asked to point out any unfamiliar words. The investigator then spelt and explained the words in question. No attempt was made to pronounce these words in order not to provide the subjects with an overt model. The subjects were then recorded reading the list twice. There was an interval of about thirty seconds between the two readings. Altogether, there were 88 and 64 stimulus items, that is two repetitions of the 44 and 34 words on the liquid and stop consonant lists respectively.

(ii) Task two: sentence reading

The second task required the subjects to read a list of sentences for which they were each given a randomized pack of cards containing the sentences. It will be recalled that these were carrier sentences which had the same test items as in the word list. The procedures followed in Task 1 were applied to Task 2.

(iii) Task three: text reading

In this task the subjects were required to read two short texts eliciting liquid and stop speech sounds. The procedures followed were basically the same as those described for Tasks 1 and 2 above. The subjects first read through the text in a familiarization session, during which they asked for the clarification of any unfamiliar items. They then read the text aloud, at a normal pace, for recording. There were two readings of each text and an interval of thirty seconds between readings.

(iv) Task 4: oral interviews.

The topics for discussion were given ahead of time in a familiarization session, during which the subjects selected two from the range provided by the investigator. Alternatively, they were allowed to suggest and discuss other topics of interest. Even though we attempted not to influence their decision, directly or indirectly, it was

necessary to restrict certain groups of learners to particular topics, in order to get them to generate as much data as possible. For example, we encouraged the low and lower intermediate learners to narrate a story or event, and to discuss their future careers or hobbies.

Each conversation lasted approximately five minutes. The investigator guided the interviews by asking questions aimed at getting the subjects to clarify, respond or to elaborate. No attempt was made to control the vocabulary used by the subjects as the aim was to gather as many liquid and final stop tokens as possible.

(v) Task 5: identification of /r/-/l/ contrasts

The aim of including an identification test in the study was to compare the subjects' perception with their production of liquid contrasts. This was a forced choice identification task which required the subjects to listen to recordings of /r/ and /l/ stimuli and label them on an answer sheet. Test recordings of native and non native speaker stimuli were presented to the subjects via a Sony MDR-24 headphone at a comfortable listening amplitude. The experimental groups were presented with stimuli produced by the Anaang speaker, and the control group with stimuli produced by the native English speaker (see Appendix 5).

It was initially intended that both recordings would be presented to non native and native speakers alike, the aim being to provide the subjects with as many

stimuli of /r/ and /l/ as possible. However, as part of the feedback received after piloting the study in Nigeria, the researcher was informed that the native speaker pronunciation was generally not distinct enough for perception. In order to establish that this was indeed the case, the same stimuli were presented to five other speakers of Anaang. We obtained the same response. On further questioning it was discovered that the main problem was with the intonation and the stress patterns which seemed to blur the actual rendition of /r/-/l/. This did not affect every word, but seemed widespread enough to warrant the elimination of native speaker stimuli from non native identification task. It is also possible that the voice pitch might have contributed to the problem. There was need therefore to control for any extraneous variables that might dramatically affect the outcome of the study.

The subjects were each given an answer sheet, A or B, to correspond with the stimuli they were presented with. The experimental groups received Answer sheet A, and the control group Answer sheet B. They were instructed to underline the words they heard, i.e. one member of each /r/-/l/ minimal pair. Each recording was presented four times consecutively, making four trials for every word.

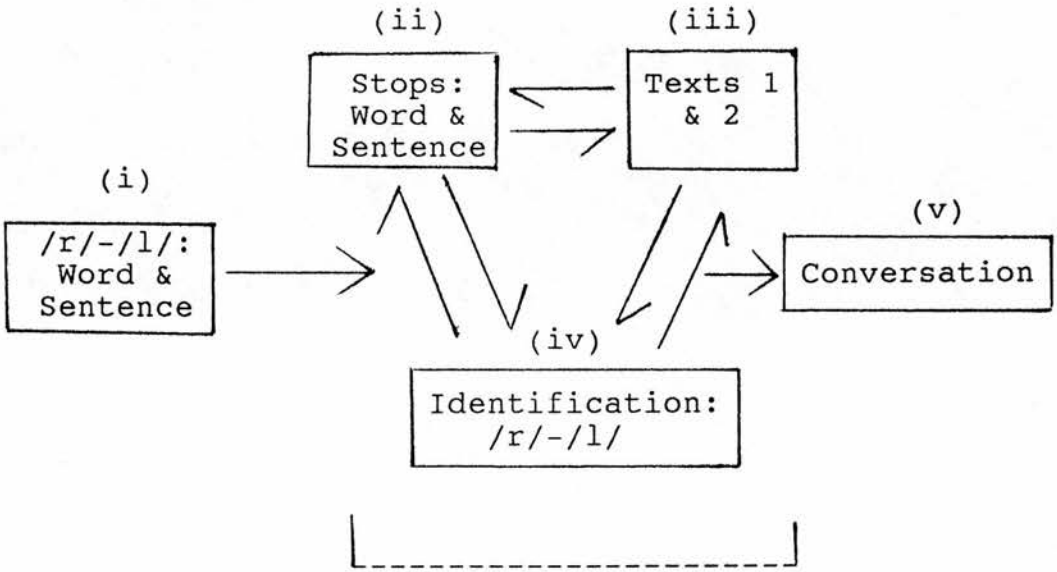
(c) Sessions

The data were collected in a single session lasting one and a half hours. Before the session began the subjects were briefed about the nature of the tasks they were going to perform and what was required of them. They were informed that the data were being collected strictly for research purposes, and would only remain accessible to persons connected with the study.

The tasks were given in a variety of sequences which were randomly assigned to each subject, although all subjects began with the /r/-/l/ word list and ended with the interview conversation. Tasks 1 and 2 in the two experiments were presented together in the sense that sentence reading immediately followed citation. That is the subjects read the entire /r/-/l/ word list first, and then read the sentences immediately after. The same procedure was applied to the stop data. With regard to text reading, Text 1 (*The Only Son*) preceded Text 2 (*Mine Boy*) each time.

The order in which the tasks were presented is illustrated below:

Figure 7.1: Order of presentation of tasks



The diagram indicates that while (i) and (v) remained constant (ii), (iii) and (iv) varied in order of presentation. The tasks were arranged in this manner to ensure that those which required a higher level of concentration (e.g. word, sentence, text and identification) were performed first while the less demanding task (conversation) was performed last. The only reason for varying (ii) to (iv) was to control for the effect of order.

The diagram yields six possible sequences as follows: A --> (i, ii, iii, iv, v), B --> (i, ii, iv, iii, v), C --> (i, iii, iv, ii, v), D --> (i, iii, ii, iv, v), E --> (i, iv, ii, iii, v) and F --> (i, iv, iii, ii, v). As already indicated above, the subjects were randomly assigned to each sequence of tasks. Details of

the manner in which they were assigned to the tasks are given in Table 7.1 below.

Table 7.1: Randomization order and assignment of subjects to tasks.

Seq-uences	A	B	C	D	E	F
SUBJECTS' IDENTIFICATION NUMBERS						
Levels						
Low	2,9,13,19	6,11,15	1,8,14,17	3,16,20	5,10,18	4,7,12
L-int	1,7,13,19	2,8,14,20	3,9,15	4,10,13	5,11,17	6,12,18
U-int	1,6,12,18	5,11,17	4,10,16	3,9,15	2,8,14,20,	7,13,19
Adv	4,7,16	8,11,20	3,13,15	2,5,14,17	6,9,18,19	1,10,13
Nss	5,20,15	1,7,9,13	2,6,11	12,14,18,19	4,3,16	10 17, 8

N.B: L-int = Lower intermediate; U-int = Upper intermediate; Adv = Advanced; Nss = Native speakers.

They performed the tasks individually before the investigator in the college Staff Room or in one of the staff offices on the university campuses. The sessions took place in the afternoons and evenings when we could ensure quiet conditions for recording. Quiet conditions were also arranged in several locations in the University of Edinburgh for recording native speaker speech. The subjects' speech was recorded on high quality

Maxell XL1-S cassette tapes using a high quality Marantz CP230 tape recorder, and an accompanying Marantz EM-8 stereo microphone which was placed at a constant distance of 15cm from the speaker's mouth.

7.3 Summary and conclusion

The main focus of the present chapter has been the description of the design of the study. The methods of investigation which include a description of the test materials, the subjects and the procedures followed have been presented in detail. As we have seen the study adopts a synchronic cross-sectional approach designed to measure performance among learners representing a range of language proficiencies. The results of the study are presented in the following chapters.