SRI LANKA CREOLE PORTUGUESE PHONOLOGY

A Thesis
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment for the Degree of
Doctor of Philosophy

by
Ian Russell Smith
May 1977

BIOGRAPHICAL SKETCH

The author was born on June 23, 1949 in Derby, England. After emigrating to Canada in 1956 his family lived in Toronto, where he graduated from Kipling Collegiate Institute in 1967. He then attended McGill University in Montreal and received a B.Sc. (Honours in Mathematics) in June 1970. In September 1970 he entered Cornell as a graduate student with a major in general linguistics and minors in South Asian linguistics and romance linguistics. He received an M.A. in linguistics in August 1973 and in September of that year went to Sri Lanka to collect data for the present volume.

Financial assistance for the author's studies at Cornell came from a bourse de doctorat from the Ministère de l'Éducation of Québec (1970-71), a tuition and fees fellowship from the Graduate School of Cornell (1971), and Canada Council doctoral fellowships (1971-76). The author also worked as a teaching assistant in French (1970-71) and linguistics (1972, 1975-76). During 1976-77 he has been an instructor in the Department of Modern Languages and Linguistics at Cornell.

The author is a member of the Linguistic Society of America, the Canadian Linguistics Association, the Indian Linguistics Association, and the Dravidian Linguistics Association. He is married to Katharine Elizabeth Ladd.

For the people of Batticaloa

ACKNOWLEDGEMENTS

I am greatly indebted to Professor James W. Gair, chairman of my special committee, whose guidance, encouragement, and friendship over the past six and a half years have been invaluable to the successful outcome of my stay at Cornell. It is unfortunate that, being away this year, he has not been able to witness the final fruition of this period.

Thanks are due also to all my teachers at Cornell and in particular to those who have served on my special committee: Professor Gerald B. Kelley, Professor Frederick B. Agard, Professor John S. Bowers, and especially Professor Larry D. King, with whom I have worked closely during the preparation of this thesis. I would also like to thank Professor Herbert L. Kufner, who guided my first efforts at teaching linguistics.

The consultants who took part in the collection of the data for this work deserve to share the credit for it, though I accept responsibility for any errors myself.

I am extremely grateful to the Canada Council, the Ministère de l'Éducation of Québec, and the Graduate School of Cornell for the financial assistance which made possible my graduate training.

My many friends in Ithaca have been a constant source of pleasure and support. They are too numerous to mention here, but I do want to express my warm appreciation for their companionship.

During the ten months my wife and I were in Batticaloa we were dependent on the good will and generosity of people who had never seen or heard of us before. To them this volume is respectfully dedicated. We shall not forget their hospitality, nor the close friends we made from among them. It is hoped that this work will at least partially repay their kindness and their trust.

The final thank you must go to my wife, Kathy, whose support, understanding and patience have at last been recompensed. Being a graduate student's spouse is a trying, at times exasperating, experience. Typing the final copy of my thesis was probably the least of her contributions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF RULES

CHAPTER II

LIST OF ABBREVIATIONS AND SYMBOLS

A	adjective
ACC	accusative
ASS	associative
BP	Batticaloa Portuguese
BT	Batticaloa Tamil
C	clitic
CNC	concord
COND	conditional
CRQ	contrastive rhetorical question intonation
CVC	Cape Verde Creole Portuguese
D	Dutch
DAT	dative
DED	Dravidian Etymological Dictionary (Burrow and Emeneau 1961)
DESC	descriptive
E	English
EMPH	emphatic
ex.	example
GEN	genitive
HON	honorific
IMP	imperative
INF	infinitive
INST	instrumental
IP	Indo-Portuguese
IT	Indian Tamil
JDBUC	Journal of the Dutch Burgher Union of Ceylon
L	a cover term for the lexical categories (N, V, etc.)
Lg.	Language, journal of the Linguistic Society of America
LOC	locative
MID	middle
N	noun, nominalizer
NEG	negative
NEGDESC	negative descriptive
NEGIMP	negative imperative
NEGPOT	negative potential
North, North IP	Northern Indo-Portuguese
OPT	optative
P	postposition
PAS	past
pej.	pejorative
PER	permissive
PFC	perfective
PL	plural

POT	potential
PPL	participle
PRES	present
QUOT	quotative
REFL	reflexive
S	Sinhala
SLCP	Sri Lanka Creole Portuguese
SP	Standard Portuguese
SPE	The Sound Pattern of English (Chomsky and Halle. 1968)
TAG	tag question marker
tns	tense
UMC	universal marking convention
V	verb
V_{int}	intransitive verb
VOLNEG	volitive negative

|	pause
-	morpheme boundary (used in the informal transcriptions of Chapter III)
#	word boundary
+	formative boundary (used in the BP underlying representations in Chapter II)
″	contrastive stress

INTRODUCTION

Sri Lanka Portuguese is a creole language spoken in Sri Lanka (formerly Ceylon). It is related to other forms of Indo-Portuguese, once spoken in various localities on the subcontinent, and eventually to other varieties of pidgin and creole Portuguese. It is important both for the study of Indo-Portuguese and for the investigation of the processes of influence involved in language contact.

Although many creoles of West Africa and the Caribbean have been well described,[1] with a few notable exceptions such as Malacca Portuguese, Mauritius French, Philippine Spanish, and Hawaiian English, creoles of other areas of the world have been neglected in recent times. Indo-Portuguese is a particular area for concern. In the late nineteenth and early twentieth centuries many varieties of this once widespread group of dialects were studied.[2] Unfortunately, the value of these works is limited by their pre-modern approach and worse, by the fact that they draw their data largely from written materials and from the casual reports of untrained contacts in the field. After a gap of more than fifty years, interest is again being shown in Indo-Portuguese.[3] Nevertheless, there is no modern description of any variety of the creole, and until one appears, work

---

[1] See Günther, B. L. da Silva and Wilson for works on West African varieties of creole Portuguese.

[2] See especially the works of Dalgado and Schuchardt.

[3] See, for example, L. Theban and M. Theban.

must be based on the earlier studies mentioned above and can hence be no more reliable than they are. Unfortunately, it appears that most Indo-Portuguese dialects are no longer available for synchronic study.[4] The Sri Lanka dialect, itself moribund, is one of the few still spoken.

Sri Lanka Portuguese is also interesting for the information it can provide on language influence. The final expulsion of Portugal from Sri Lanka in 1658 effectively removed Standard Portuguese as a model for the creole to emulate. Most of the world's creole languages have remained in contact with the standard languages on which they are based. The greater prestige of the standard usually leads to attempts by creole speakers to imitate aspects of its structure. The result is that it is often difficult to determine the original nature of the creole. Sri Lanka Portuguese is an exception to this pattern: for the last three centuries only the few educated creole speakers have had any knowledge of Standard Portuguese. Thus the creole has been able to develop independently from its mother language. On the other hand, it has been profoundly influenced by Sinhala and Tamil, the indigenous languages of Sri Lanka, with which it has been in intimate contact ever since its establishment in the island. Dutch, and more recently English, have also had some influence on the lexicon and phonemic inventory, but none, as far as I can tell, on other aspects of the grammar.

---

[4] L. Theban. Personal communication.

This work is based on data gathered during ten months' field research in 1973-4 on the Sri Lanka Portuguese spoken by the 'Burgher' community of Batticaloa, a provincial town on the east coast of the island.

The data gathered on Batticaloa Portuguese was of two types: first, material obtained during formal elicitation sessions; and second, unrehearsed conversations and monologues recorded live and later transcribed. In addition, a few folk songs were also recorded and transcribed. The elicited data are the more systematic, but because the language itself is the focus of attention in the elicitation situation, they could be subject to various kinds of distortion. As a check against this, at the end of each session (usually about an hour in length) a rapid re-elicitation of the session's data was recorded. The taped data was later compared to the field notes and differences noted. The conversations and monologues entail a completely natural use of language[5] and thus provide a further check on the elicited material. In the field they also furnished new material for investigation in the elicitation sessions.

All the Batticaloa Burghers speak Tamil, the dominant language of the area. Although the main focus of field work was the creole, some work was also done on Batticaloa

---

[5] I was present during the taping of these, and occasionally took part in them. This practice of course means that my and the tape recorder's presence must be counted as parameters of the speech situation. However, this caused no apparent inhibition of spontaneity or other register effects.

Tamil, which differs from both continental Tamil and Jaffna Tamil. In addition to the usual elicitation sessions, some of the conversations recorded in Batticaloa Portuguese were translated into Batticaloa Tamil by a bilingual mother-tongue speaker of the creole. The Tamil versions were subsequently verified with a mother-tongue speaker of Batticaloa Tamil.

The creole corpus consists of about 6,000 entries (an 'entry' may be a word, a phrase, or a sentence). On tape, the elicited material occupies just under seven hours and the spontaneous material just over three hours, about half an hour of which was translated into Tamil (roughly 450 entries). The elicited Tamil corpus contains approximately 2,000 entries.

The consultants used in the study are listed below. Not included are many others to whom I am indebted for participating in the various recorded conversations or otherwise giving informal help.

Ronald Rosairo

Born: Batticaloa, 1906

Mother tongue: Sri Lanka Portuguese

Other languages spoken: English, Tamil, Sinhala

Literacy: English

Education: seventh standard (English medium)

Occupation: Planter, owner of an auto and bicycle tire dealership, Justice of the Peace

Mr. Rosairo was my chief consultant for Batticaloa Portuguese. He takes considerable pride in the linguistic and cultural heritage of the Burgher community and is president of the Catholic Burgher Union. He uses the creole as the language of the home and of work and is recognized as a highly proficient speaker.

Richard Starack

Born: Batticaloa, 1928

Mother tongue: Sri Lanka Portuguese

Other languages spoken: Tamil, Sinhala

Literacy: Tamil, Sinhala

Education: seventh standard (Tamil medium)

Occupation: machine operator

Mr. Starack's family moved to Valaichchenai when he was 15, and at age 25 he went from there to Amparai (in the Sinhala speaking area) for work. He returned to Batticaloa at age 28. He speaks creole at home. He was especially helpful in relating creole and Tamil structures.

Father Norbert Ockersz S.J.

Born:

Mother tongue: Sri Lanka Portuguese

Other languages spoken: Tamil, English, some Sinhala and Latin

Literacy: Tamil, English, Latin

Education: SSC and seminary training (seven years)

Occupation: Roman Catholic priest

Father Norbert went to India in 1946 for education. Since his return in 1953, he has served in various parishes outside Batticaloa. In 1975 he worked for several months in Hamilton, Ontario, where I interviewed him. Spending most of his time away from Batticaloa, he speaks creole less frequently than my other consultants; however, he is still fluent and was able to provide some extremely useful additional data.

Jothy Jeevaratnam

Born: Batticaloa, 1950

Mother tongue: Tamil

Other languages spoken: English, some Sinhala

Literacy: Tamil, English

Education: GCE (English medium)

Caste: Vellāḷar

Occupation: Rock drummer, lawyer's assistant

Mr. Jeevaratnam was my main consultant for Tamil

This work is divided into three chapters. Chapter I provides a sketch of the (extra-linguistic) historical and socio-cultural setting of Batticaloa Portuguese. A synchronic description of the phonology may be found in Chapter II. Chapter III presents an assessment of indigenous influence

on the creole's phonology, which leads to a consideration of what is entailed in inter-language influence

CHAPTER I: THE SETTING

In this chapter, we first give a brief overview of Sri Lanka and its history, and with this as background, attempt to sketch in some of the details of the social history of creole Portuguese in the island. A full investigation of this matter would involve thorough study of primary sources and is beyond the scope of this work. What is offered here relies entirely on published material and contains many gaps in information and much conjecture. Nevertheless, I believe it is, in the main, an accurate account. This will be followed by a short description of Batticaloa.

1 Sri Lanka

Sri Lanka is an independent island nation of 25,332 square miles, lying some 30 miles from the southeastern coast of India. It has a mountainous south-central region with 15 peaks over 6,000 feet; this is surrounded by a broad coastal plain of less than 500 feet in altitude. The population of 13 million is predominantly (77%[1]) rural, and is most concentrated in the southwest of the island, where the capital, Colombo, is also located.

Several linguistic, ethnic, religious and caste groupings exist. The major languages of the island are Sinhala, an Indo-Aryan language spoken over all the island, except

---

[1]Population statistics quoted in this paragraph and the next are from the 1971 census.

the northern tip and a strip running down the east coast, and Tamil, a Dravidian language also found in southern India and Malaysia, which is spoken in the north and east. Sinhala is spoken by the predominantly Buddhist Sinhalese, who form 72% of the population. Tamil is spoken by the largely Hindu 'Tamils' and the Muslim 'Moors', respectively 20.5% and 6.7% of the population. Christians (7.7%) can be found among both the Sinhalese and the Tamils. Sri Lanka Portuguese is spoken by an unknown, but certainly extremely small, number of people.[2] All known speakers are members of the 'Burgher' community: descendants of the Portuguese and Dutch who founded families in Lanka. The Burghers account for 0.3% of the population.

The history of Sri Lanka relevant to the creole begins in 1517, when, attracted by the island's spices and strategic position (midway between their holdings on the west coast of India and Malacca), the Portuguese sent an expedition from Goa to establish a fortified trading post at Colombo. By exploiting the unstable political situation in the island at the time, the Portuguese soon won for themselves the position of guardians of the nominal monarch of southern Lanka, Dharma Prakrama Bahu, who held court at Kotte, near Colombo. In 1597 Don Juan Dharmapala, the last king of Kotte, died childless, willing his realm to Phillip II, king of Portugal. With the annexation of Jaffna in

---

[2]Unfortunately, no census taken in the island has ever included figures on the language.

1617, Portuguese authority extended over the whole of the lowland zone.

The Portuguese did not have an easy time in Lanka, however: the Sinhalese of the interior waged continual war against them, first under Maya Dunnai (nephew of Dharma Prakrama Bahu) and his son Raja Sinha. The leadership then passed to the kings of Kandy, as the still independent highlands (Sinhala /kandə uḍə raṭə/ 'country on the hill') became known. Colombo and Kotte were besieged so many times that in 1563 the latter was abandoned. The Portuguese erected forts at Colombo, Point de Galle, Negombo, Matura, and eventually at Batticaloa and Trincomalee; and retained up to 10,000[3] soldiers, of whom only 1,000 were Europeans. The expenditure required for these protection measures was so great that it consistently outstripped the value of the colony's trade. The Portuguese were more successful in their efforts to spread the Roman Catholic faith, which has a substantial number of adherents even today.

The Dutch had had contact with the Kandyan court as early as 1602, but it was not until 1632 that the recently enthroned Raja Sinha II invited Dutch cooperation in expelling the Portuguese from the island. There followed a long

---

[3]The number of Europeans is well documented in Portuguese records, but the precise number of local men, known as lascarins, serving with the Portuguese is not known. Tennent's claim of an army of 'upwards of 20,000 men' (1859 vol. II p. 26) is doubtless fanciful. C. R. De Silva posits about 8,000 for the 1630's and 'a few thousand more' for the 1620's (p. 188). But earlier in the same work (p. 91-2) the figure of 5,000-6,000 is given for the 1620's.

period of conflict, punctuated by truce only from 1645-1650. The Dutch took Batticaloa in 1638, but Colombo did not capitulate until 1656; conquest was complete with the fall of Mannar and Jaffna two years later.

When Raja Sinha II realized that, counter to his expectations, the Dutch were not about to deliver their new conquests to him, the alliance quickly dissolved into enmity. However, as the Dutch East India Company, which administered the island, was primarily interested in the profits to be derived from commerce, it resisted as long as possible engaging in costly military operations against Kandy. During the Dutch reign, the years 1761-1766 mark the only period of outright war.

Meanwhile, Britain had been casting a covetous eye toward the island. Early missions to the Kandyan court in 1763 and 1782 amounted to nothing, but in 1795, when France invaded the Netherlands, the British took advantage of the weakness of the Dutch and sent an expedition to Lanka from Madras. With Kandyan support the British easily took Trincomalee, Batticaloa, and Jaffna in the same year. Colombo and the remaining Dutch possessions surrendered in 1796. Holland formally ceded the territory to Britain in 1802 by the Treaty of Amiens.

Relations between the British and the Kandyans deteriorated steadily. Their decline was aggravated by the intrigues of two successive Adigars (chief ministers) of the Kandyan court, who sought to have the king deposed and

to ascend the throne themselves. An invasion of Kandy in 1803 was repelled, but in 1815 a more successful expedition finally quelled the highland kingdom, and for the first time the whole of Lanka was brought under the domination of a foreign colonial power.

The island regained its independence in 1948 as the Dominion of Ceylon. On May 22, 1972, a new constitution was promulgated and the Republic of Sri Lanka was established.

2 External History of Sri Lanka Portuguese

The Portuguese were established for twenty years in Goa before coming to Lanka in 1517. It is likely that by then a distinct Indic variety of pidgin Portuguese had begun to develop, and that this was used as the basis for communication with the new territory's inhabitants. Although the political occurrences of the Portuguese era are well documented, very little information is available on the linguistic situation. It is clear, however, that by the early 17th century a Portuguese-based pidgin was in use in the Portuguese-controlled littoral, and was not unknown in the kingdom of Kandy by virtue of its frequent dealings with outsiders. Moreover, a creole community had sprung up. There were probably two groups of creole speakers: the so-called Topazes (Tupasses, etc.), 'dark skinned or half-cast claimants of Portuguese descent, and Christian profession',[4]

[4] Hobson-Jobson. It is possible that in Sri Lanka it was

and Kaffirs (Caffres, etc.), or East Africans.

The Topazes, children of local or half-caste mothers and Portuguese or half-caste fathers, would have been exposed to pidgin/creole Portuguese at home.[5] As the above definition implies, they identified with Portuguese, a not unnatural phenomenon, considering that the Portuguese were at the apex of the social order, though they must have had local family ties as well. They fought in the Portuguese army, where they served as reserves, much as the Portuguese Casados (married men who had come with their Portuguese wives as settlers to Lanka). In 1593 there were 350 Topaz troops.[6]

Black slaves were brought to Lanka by the Portuguese from East Africa. There is no doubt that these people would have spoken pidgin Portuguese, and it is likely that many were native creole speakers, though perhaps not of the Sri Lanka variety. As domestic servants they would have introduced the children of Casados to the pidgin/creole at an early age. Kaffirs also served in the army, though according to Silva they were not used in this capacity in large numbers until 1631, and by 1634 there were only 280

---

only those with quite dark skin who were called Topazes. This is irrelevant to our point, which is that intermarriage was commonplace.

[5]This is an oversimplification. There was undoubtedly much variability in the range of speech varieties controlled by both parents, depending on their previous experience with the pidgin/creole in Lanka. The situation would also have changed considerably during the near century and a half of Portuguese domination.

[6]Abeysinghe p. 188 n. 14.

fighting men left.[7] Whether there were also children of chance unions of Portuguese or Topazes with Kaffirs, and to what group these would have belonged is not clear.

By the time the Dutch arrived on the scene, there was a reservoir of mother-tongue creole speakers established in Lanka, and their language was in widespread use as a lingua franca. In 1663, Van Goens, the new Dutch governor writing from Colombo to the Seventeen (the directors of the Dutch East India Company in Europe) shortly after taking up his post, reported that

> Colombo and Galle swarm with evil-disposed Portuguese of the lowest kind who refuse to work and, increasing fast, like weeds are a great danger to the community ... (Pieris 1929, pp. 280-1)

Toussaint writes:

> The [pidgin/creole] Portuguese language was well established in Ceylon at the time the Dutch took possession of the maritime provinces. It was freely spoken by the Sinhalese, and it was the language used in all interviews with the Kandyan Court. (1942, p. 39)

With the departure of the Portuguese, the standard language of Portugal played no further role in the development of the creole in Sri Lanka: it is unlikely that a significant number of the Europeans in the service of the Dutch East India Company would have spoken Standard Portuguese. It is true that some of the widows and fatherless daughters of Casados killed in battle were not deported along with

---

[7] P. 188.

the rest of the Portuguese, but retained by the Dutch as wives for their own men. However, as their numbers were very small, their effect can be discounted. Later speakers of Standard Portuguese who visited the island as missionaries, etc. could well have had some effect on the way the creole was written, but at the colloquial level, again because of their numbers, their influence could not have been significant.

As had been mentioned, trade was the primary concern of the Dutch. Thus even though their official policy was to eradicate the language and religion left by the Portuguese and to supplant them with Dutch and the Dutch Reformed Church,[8] in practice Roman Catholicism proved capable of withstanding the most severe repressive measures, and pidgin/creole Portuguese was not only tolerated, but extensively used by the Dutch themselves.[9] In his journal, Christopher Schweitzer, a German serving with the Company in Sri Lanka from 1676 to 1682, records that as one of his first acts after arriving in the island he

> learned the Cingulaish and [pidgin/creole] Portuguese languages, finding them not only useful, but necessary to those who are to stay there. (Hulugalle p. 98)

During the early years of Dutch rule, intermarriage was

---

[8] Cf. Tennent 1850 pp. 37-76 and 1859 vol. II pp. 57, 69, 70, et passim.

[9] Valkhoff argues that creole Portuguese was to be found in other Dutch possessions of the 16th and 17th centuries, including Batavia and the Cape, which had never been in Portuguese hands. (pp. 146-7 et passim).

common, as it had been under the Portuguese.[10] Schweitzer has many references to the state of affairs:

> The Europeans here ['Anguratct', i.e. Anguruwatota], about Forty to Fifty in number, live in the following manner ... Out of the Fort ev'ry one hath his Cingulaish Woman to provide and dress his Victuals. If a Woman at any time is brought to bed of a white Child, it is a matter of great Joy to her Parents and Friends, ... (Raven-Hart p. 58)

> The Fort [at 'Manara', i.e. Mannar] ... is kept by near 100 Dutch Soldiers. Each Soldier keeps a Boy to clean his Arms and carry them on the march for him and a Woman to look after his Meat and serve him. (Raven-Hart p. 62)

> And the Marriages [of the Company's men] in Columbo are performed as follows ...
>
> The Ceremony is usually held in Portuguese. ...
>
> Any European in this Island under Dutch rule may marry any Woman he pleases, of whatever Nation she may be, if only she is Christen'd and has a Certificate from the Minister. (Raven-Hart p. 76)

> The Women [in such unions] lead very lazy Lives; ... They leave all the Care to the Man, that he shall get the Mony for the House; and they generally have two Slaves. (Raven-Hart p. 77)

In these marriages, creole Portuguese was the medium of communication, as Anthonisz relates:

> The early Dutch settlers then who came to live with and among these people had to learn their language. It was the language used for all domestic and ordinary purposes ... Dutch children born in Ceylon learnt it from their nurses, and used it as a home language with greater freedom than their mother tongue. It became in fact to a great extent the domestic language in Dutch households. (pp. 30-31)

---

[10] Cf. Pieris (1918) p. 6 et passim, Arasaratnam p. 206.

By the turn of the century, however, according to Pieris,

> intermarriage was officially discouraged, for it was thought that the two generations which had come into existence since the occupation of Galle in 1640, had produced enough daughters with European fathers to satisfy the requirements of the Hollanders. (1918, pp. 154-5)

By attempting to limit the choice of possible brides in this way, the Dutch subdivided the creole community, creating a new Netherlands-oriented elite. Even so, creole Portuguese continued as the home language of this group. We again refer to Anthonisz:

> While Dutch became the spoken and written language of polite society, the language used in the office, the platform and the pulpit, a form of Portuguese, in which a large proportion of Dutch words had found admission ... was used as the medium of conversation in familiar gossip and in the most intimate relations ...
>
> This was the state of things that existed here when the English arrived in the last years of the eighteenth century. (p. 31)

The situation in Sri Lanka would thus appear to parallel that in Surinam, where Dutch failed to supplant the previously established English-based creoles.

The Dutch-oriented creoles, following their overlords,[11] held the 'Portuguese' in contempt. The division in the creole community must have come to follow both color and socio-economic differences, if indeed it did not reflect earlier stratification along these lines. The creoles were essentially an urban community, having neither land nor a

[11] Cf. the quote from Van Goens above.

knowledge of agriculture; with the Dutch and their camp-followers monopolizing the administration and trade, the 'Portuguese' would have had little means of supporting themselves. Also the continuous influx of European men needing to be provided with wives would certainly have tended to produce a lighter skin color in the Dutchified creoles if they had not been drawn from lighter stock.

The Dutch continued the practice of bringing Africans to Lanka, thus contributing to the 'Kaffir' community. They also introduced people from the islands of the Malay archipelago, chiefly for military service, at which they apparently excelled. Denham notes that at the end of their reign, the Dutch had more than 15 companies of Malays in their service.[12] These people, apparently remained Muslims and did not contribute to the wholly Christian creole community. They did, however, intermarry with the Sinhalese and Moors.[13]

From the beginning of the British period more frequent references to the creole can be found. During the nineteenth century numerous general works on the island were published. One of the earliest of these, Cordiner's Description of Ceylon, describes the situation soon after the advent of the British. Here we find described for the first time in print the three sections of the creole speaking community, which I shall henceforth refer to by

---

[12] P. 238.

[13] Cf. Denham p. 238.

their modern names: Dutch Burghers, Portuguese Burghers, and Kaffirs:

> The Dutch inhabitants in Ceylon are about nine hundred in number, and, excepting a few families, are reduced to circumstances of great indigence: but by rigid and meritorious economy, and some of the lesser labours of industry, they maintain an appearance, in the eyes of the world, sometimes affluent and gay, always decent and respectable.
>
> They are chiefly composed of officers (prisoners of of war), with their families, and widows and daughters of deceased civil and military servants of the Dutch East-India Company. The greater part of them are proprietors of houses, which they let, with considerable advantage, to the English inhabitants. (vol. I, p. 87)

> The Dutch settlers in Ceylon use no other servants but slaves, a family of whom always composes part of their household. (vol. I, p. 81)

> There is still a large body of inhabitants at Columbo and the other settlements in Ceylon, known by the name of Portuguese. They probably number to the amount of five thousand; they are, however, completely degenerated and exhibit complexions of a blacker hue than any of the original natives. Yet they retain a considerable portion of the pride of their ancestors: wear the European dress: profess the religion of the church of Rome; and think themselves far superior to the lower classes of the Cingalese. The greater part of them were admitted, by the Dutch, to all the priviledges of citizens, under the denomination of burghers. A corruption of their original language is still spoken over all the sea coasts. It is very easily learned, and proves of great utility to a traveller, who has not time to study the more difficult dialects of the natives. (vol. I, pp. 88-9)

> The usual garrison in Columbo is one regiment of British soldiers, one regiment of Sepoys, or Malays, one company of gun-lascars, and a small corps of native pioneers. To the above establishment, a body of seven hundred Caffrees has been lately added, and formed into a regiment. Many of them were slaves at the Portuguese settlement of Goa on the coast of Malabar, where they were purchased by our government ... They are all

> nominally Roman Catholic Christians, and certainly know nothing of any other religion. On Sundays they march to the Portuguese church ... Many of them have brought wives and children with them. (vol. I, pp. 65-6)

The British themselves did not marry into the creole community, but remained aloof in their social relations.[14] And, although in 1806 Sir Thomas Maitland, the governor, enacted a regulation 'for taking off the restraints which were imposed upon Roman Catholics of the island by the late Dutch Government,'[15] it is clear from Cordiner above, as well as Selkirk below, that the British held the Portuguese Burghers in the same low esteem as had the Dutch.

Both Dutch and creole Portuguese soon began to yield to English. The position of Dutch was the more precarious - it had been the high language of the elite, but the Dutch Burghers did not for the most part speak it at home. Now there was a new elite and their language usurped all the functions formerly reserved for Dutch, except for a time its rôle in the Dutch Reformed Church. The bulk of the Dutch judicial and administrative staff were allowed to remain in their positions under the British, and a knowledge of English was thus essential to them:

> Every opportunity was sought for acquiring a knowledge of the new language; and, it is easy to explain how, in the scramble for Government employment, the cultivation of Dutch was neglected. (Anthonisz p. 34)

---

[14] Cf. Cordiner vol. I, p. 76.

[15] Full text appears in Tennent 1850 p. 75.

The British were certainly not keen to see Dutch perpetuated: in 1801 it was abolished from the courts of law and successive administrators sought to discourage its being taught in the schools.

No such pressure was put on the creole, and for some time it apparently continued to be used even in the courts.[16]

> It continued to be the colloquial language in the households of the best Dutch families. The introduction of English had in no way affected it, as it had never [been] brought into conflict with it. (Anthonisz p. 35)

In the early 1800's the Weslyan Church became active in the island and found the creole to be a useful tool. Its missionaries produced several dictionaries, phrase books etc.[17] on what by then was called 'Ceylon-Portuguese'. Religious literature also began to appear.

Writing in 1844, Selkirk, author of another general work, reports:

> The PORTUGUESE are to be found in every large town in all parts of the island, as well as in many of the smaller ones. They abound, however, most in Colombo and Jaffna ... In colour they resemble neither the Singhalese nor the Tamulians, some are of a duller black then either, and others of a sickly yellow. They keep up the European dress ... Without capital to embark in trade, or to purchase government lands, and too pround to rent and cultivate an estate, even if they had capital to purchase it, and not sufficiently trustworthy to be employed by government, or in any responsible situation among merchants, they pass their time in idleness, and filth,

---

[16] Cf. Vos p. 134.

[17] See the works of Callaway and Fox in the bibliography. A fuller listing can be found in Reinecke et al.

> and sin. Those who do work are generally tailors, shoemakes, bazarkeepers, etc. Learning is at a low ebb among them; and the only books in their language, are parts of the Bible, and the whole Book of Common Prayer, both of which have been translated for them by missionaries within the last twenty-five years ... they have of late begun to pay some attention to the English language. (pp. 69-70)

> The DUTCH, who, together with the Portuguese, are called the "Burghers," form but a small portion of the inhabitants of Ceylon ... Their own language is not much used among them except it be among the old ladies. The common language used in their families is the Portuguese. Many of them understand and converse in English. They have filled situations of importance and respectability under the English government, to whom their services have been of great value. (p. 72)

> The Caffres ... speak the Portuguese language and are all, or most of them of the Roman Catholic religion ... they have the thick lips, high cheek bones, and curly hair, of the natives of Caffraria ... They seldom intermarry with persons of other tribes. (p. 75)

Concerning this last group Marshall writes in 1846:

> they have never kept up their numbers, and have nearly disappeared. (p. 18)

Tennent (1850) gives a more humane description of the Portuguese Burghers' situation:

> The latter [i.e. the Dutch Burghers], from their education and capabilities, aspire to offices of rank and responsibility; whilst the Portuguese section of the "burghers" are contented to fill the humbler occupation of tradesmen and artisans. The distinction, thus perpetuated, till it has almost become a characteristic of race, may be traced to the despotic and fanatical policy of the Dutch Government in Ceylon, which, from the commencement of its rule, subjected the Portuguese, on the plea of their being Roman Catholics, to social degradation, excluded them from every office of emolument, and effectually shut them out of every pursuit of industry or path to distinction. (p. 72)

The same author, writing again in 1859, observes:

> Already the language of the Dutch, which they sought to extend by penal enactments, has ceased to be spoken even by their direct descendants, whilst a corrupted Portuguese is to the present day the vernacular of the middle classes in every town of importance. (vol. II, p. 70)

Shortly thereafter, though, the language seems to have entered a period of decline, gradually yielding its function as a lingua franca to English. In 1880, Coehlo reports that[18]

> A missionary who was in the island wrote on the date of the 13th of November of 1875, that Indo-Portuguese is almost exclusively the language of the descendants of the Portuguese and Dutch who settled in the island, that the language is not considered important by missionaries as a medium of instruction, insomuch that those who use it also speak another language, that the dialect is in extreme decline and that during the course of another generation it will become totally extinguished. (pp. 157-8)

During the nineteenth century the Portuguese Burghers fared little better under the British than they had under the Dutch. After nearly a century of British rule Cumming paints the following picture:

> The [evangelistic] work among the latter [i.e. the Portuguese Burghers] is most discouraging, the majority being so steeped in hopeless poverty that their life seems to have lost all spring ... Their own language is a very debased Portuguese. Of course the well-to-do Dutch Burghers form a large and very important class of the community. As may be guessed by a glance round any of the churches one may chance to enter, they fill all sorts of responsible positions, but the Portuguese seem

---

[18] I have translated this and all other quotations from sources not in English.

> never to have got over the crushing oppression to which their ancestors were subjected by the Dutch, and to this day few rise high in the social scale. (vol. II, p. 425)

In 1900 Dalgado, a Konkani Jesuit, who was Vicar-General of Ceylon in 1886, produced an extensive work on the creole, which comprises a grammar, anthology and vocabulary. Contrary to Coelho's indications, he writes:

> Even at the present time, despite a century of the English domination with its powerful colonializing administration and extensive instruction of its own language, the Indo-Portuguese of Ceylon or português basso (low), as the dialect is commonly called there in contrast to português alto (pure), shows no signs of dying, but on the contrary, promises a long and robust life, adjusting itself, of course to the circumstances, and being subjected to incessant dialectal evolution ...
>
> it is spoken as a mother tongue by nearly all the Eurasians or Burghers, and learned for commercial domestic and religious use by many Europeans and natives. (pp. xxi-xxii)

This is probably an over-optimistic picture. Certainly even the Dutch Burghers were beginning to give up the low-prestige creole in favor of English. In 1908 Anthonisz, himself a Dutch Burgher, writes of the creole:

> The Dutch Burghers of the present day, with the aid of improved systems of education, have completely weaned their children from any inclination to speak it; and, to the younger generation of the Dutch Burghers, Portuguese, in any form is a wholly unknown tongue. It still prevails in the island only among the people described by the Dutch as ambachtslieden, a class which, by the literal translation of this term, into English has come to be known amongst us as "mechanics." (p. 35)

Against this, in 1911 Vasconcellos claims that 10% of

the population of Colombo spoke the language.[19] Anthonisz, with the opportunity for firsthand observation is probably the more accurate, though he is clearly trying to downplay the creole's importance. His claim that the Dutch Burghers had given up the creole is probably valid only for the well-to-do Burghers of the Dutch Burgher Union, who were primarily found concentrated in and around Colombo.[20] As the Batticaloa evidence will show, many less priviledged Dutch Burghers retained the language and eventually lost their separate identity and merged with the Portuguese Burgher community. The number of Burghers in Sri Lanka in 1911 was reported by the census of that year as 26,663.[21]

The Kaffirs in Lanka have declined greatly in numbers. Brohier reports visiting a Kaffir 'colony' called Sellan Kandal near Puttalam around 1920:[22]

> I was definitely struck by the preponderance of characteristics of definite Negroid origins - frizzly curled, or wooly heads of hair, and skins like polished ebony. There were however, few survivors of the original Kaffir stock for they had intermarried freely with the mixed Portuguese descendants in the Island ... The majority of the colonists spoke a very corrupt form of Portuguese, and most of them apparently were Roman Catholic adherents. (p. 25)

Sri Lanka Portuguese largely disappeared from view

---

[19] P. 316.

[20] Of the 300 odd members listed in 1908 (JDBUC vol. I, no. 1, pp. 61-7) 75% lived in Colombo. No other area had more than 5% of the membership.

[21] This figure also includes other Eurasians.

[22] See also Raghavan, p. 200.

during the middle part of this century, though a non-linguistic article by Vos in 1950 does at least confirm its existence. In recent years Dr. D. E. Hettiaratchi, professor emeritus of the University of Sri Lanka, Peridéniya, and editor-in-chief of the Sinhala Encyclopaedia, has taken a strong interest in the creole and actively sought out speakers. In an article in 1969 he reports:

> There are several scattered groups of families earning their living as shoe-makers, carpenters, blacksmiths, motor mechanics, and other craftsmen in towns like Colombo, Kandy, Trincomalee, Galle, and Akkaraipattu, and there are also small communities of them at Batticaloa and Chilaw. (p. 774)

He also mentions a Kaffir village near Kata Oya, near Puttalam, some of whose residents claim to be descendants of a regiment from Madagascar, which came to Sri Lanka during the Boer War:

> of their descendants today, about twenty families are said to be at Srambiadiya, three miles from Puttalam, about fifteen families at Kundasale near Kandy, several families at Kurunegala, and several at Nuwara Eliya in tea plantations. Only the elderly people among them have kept up their Portuguese speech. (p. 747)

3 Batticaloa

When the Dutch Admiral Joris von Spilbergen landed near Batticaloa in 1602 and proceeded on to Kandy, Portuguese authority in the east of the island was largely theoretical; in practice the many petty princes of the region were sympathetic to the Kandyan cause. Thus the Dutch were able

to enter and leave at will. The Portuguese did finally establish a permanent presence in Batticaloa in 1627 by building a fort on Puliyantivu, an island in the Batticaloa lagoon. This was captured by the Dutch in 1638. Vice-Admiral Willem Jacobsz Coster directed the assault. The following is excerpted from his report to Governor-General Antonio van Diemen at Batavia, dated June 4, 1638:[23]

> On the 18th [April] ... our batteries played merrily on the fort ... until the enemy put up a flag of truce and surrendered the fort to us ... The conditions were that they should go to Negapatam and the blacks do homage to the King. There were about 700 in all, among whom were 50 Portuguese and mestices, the rest being blacks, women and children.

The Dutch repaired the fort and maintained a garrison there for some time. But since no articles of trade were available there to offset the cost, they regarded Batticaloa as a liability; accordingly, in 1643 their occupation was brought to an end. The fort was given over to the King of Kandy and was demolished by him. The Dutch reoccupied Batticaloa sometime between 1646 and 1665, intending to use it as a base for trade in arecanuts, elephants, and other agricultural products (rice was in surplus in the region) as well as for the collection of tolls and taxes. They were hampered for a long period by the fact that their possessions were limited to Puliyantivu. By the treaty of 1766 Kandy ceded all coastal territory to the

---

[23] Ceylon Literary Register II (August 5, 1887). Quoted in Toussaint 1929.

Dutch, and this addition made the station more viable economically.

Unfortunately, due to the loss of the Batticaloa records, many details of the Dutch occupation are not available. Regarding those stationed there late in the Dutch period, Toussaint writes:

> The Dutch depended a good deal for their labour on slaves. The Company's slaves in Batticaloa in the year 1794 numbered 92 viz., 26 men, 32 women, and 34 children. (1931 p. 101)

> the Dutch population in Batticaloa consisted exclusively of officers of the Dutch East India Company and there were no private citizens (1931 p. 102)

> The number of Dutch people could not have been much more than thirty, and the number of Dutch women must have been very small indeed if there were any at all. (1931 p. 103)

In 1795 Batticaloa was occupied by the British. Even though a district administrative center during the British and post-independence periods, Batticaloa has always been an isolated outpost. Tennent's comments are as apt today as they were in 1859:

> Far less frequented by Singhalese and Europeans than any other portion of Ceylon, the Eastern Province has retained many ancient habits, and presents more frequent instances of curious social peculiarities than are to be noticed in the rest of the island. (vol. I, p. 458)

This isolation has undoubtedly been a factor in the preservation of the creole. Not surprisingly, very little information about the town's history can be gleaned from the

literature.

An on-the-ground study using local records would be necessary to obtain any really useful information. Particularly valuable would be the records of the Catholic Church. I was puzzled, for example, to find that though the Burghers are all Catholics and identify themselves as 'Portuguese', a great many of them have Dutch surnames. I was told that at one time there were also Anglican and Wesleyan Burghers in Batticaloa, but that due to the dwindling number of Burghers in their churches, they all converted to Catholicism. The Catholic church maintains records of births and marriages for the Burgher 'caste'. Early records, which go well back into the last century, contain only Portuguese names. In the latter part of the century, though, there was a sudden influx of Dutch names. The interpretation which immediately suggests itself is that unlike their well-to-do counterparts in Columbo, the Dutch Burghers in Batticaloa never gave up the creole, though the distinction between Portuguese and Dutch Burghers was maintained well into the British period, in part through a difference in religion and lack of intermarriage. Eventually, the two groups must have found they had more in common than they had differences. This would have especially been true when the creole lost its status as a lingua franca.[24]

---

[24]Another possible interpretation might be that the Dutch Burghers were an immigrant group.

Present day Batticaloa is a medium-sized coastal town (in population, it ranks tenth among Srï Lanka's urban areas) and administrative center for the district of the same name. Like the rest of the Eastern Province, Batticaloa District (area 1,017 sq. mi.) has a heavily populated coastal area and a sparsely populated interior. The mainstays of its economy are agriculture, fishing, and the paper mills at Valaichchenai, the region's only major industry. The area is culturally quite diverse; in Table 1 comparative figures for ethnic group membership and religion are given for the district and town as well as for Sri Lanka as a whole. Tamil speakers are in the overwhelming majority and there is no reason to doubt that such was not the case throughout the colonial period.[25] English and Sinhala are also in currency, particularly in the urban areas.

As can be seen from Table 1, the Burghers themselves are concentrated in Batticaloa Town. I was told that in the early part of this century most Burghers lived close to the center of town. Many have since moved to outlying areas, especially Dutch Bar and Cinna Uppōṭai.[26] The Burghers are for the most part of low economic status, traditional occupations being carpenter, blacksmith, bicycle (and more recently motorcycle and car) mechanic, and laborer. Many

---

[25]Cf. Tennent above.

[26]Burghers who have gravitated to the large Batticaloa community from smaller villages and towns of the region have also added to the population of these areas. (Dennis McGilvray - personal communication)

Table 1.[+] Population of Sri Lanka, Batticaloa District and Batticaloa Town by Ethnic Group and Religion

	Sri Lanka		Batticaloa Dt.		Batticaloa Town	
	number '000	%	number	%	number	%
All groups	12,711	100	258,104	100	36,761	100
i) Ethnic Group						
Sinhalese	9,147	72.0	11,282	4.4	2,534	6.9
Tamils	2,611	20.5	182,661	70.8	29,685	80.0
Moors*	853	6.7	61,765	23.9	2,876	7.8
Burghers**	44	.3	2,255	.9	1,676	4.6
Malays	42	.3	59	.02	17	.05
Others	14	.1	82	.03		
ii) Religion						
Buddhists	8,568	67.4	9,624	3.7	***	***
Hindus	2,239	17.6	167,597	64.9	***	***
Christians	987	7.7	18,116	7.0	***	***
Muslims	910	7.2	62,519	24.2	***	***
Others	8	.1	48	.02	***	***

[+]Adapted from Statistical Pocket Book of Sri Lanka and an untitled handbook prepared by the Batticaloa Kachcheri. Figures are for 1971.

*'Moors' are predominantly Tamil-speaking Muslims.

**Includes non-Burgher Eurasians.

***No figures available.

are unemployed. However, a significant minority have managed to attain higher status by obtaining clerical and administrative positions in the state corporations and private enterprise, or by starting their own businesses.

In physical appearance they are quite variable. Some have skin as dark as that of the Sinhalese and Tamils; others are light as a European. A few have curly hair and protruding lips, due probably to an African factor in their ancestry.

Culturally, while maintaining its group distinctiveness, the Burgher community has become largely assimilated to the matrix culture. The group is largely endogamous and functions as a caste-like unit; cross-cousin marriages are traditionally highly valued; group identity is furnished by dress, certain foods, customs associated with marriage, and, of course, language.

All of the Burghers in Batticaloa speak Tamil, many of them better than they do the creole. Some also speak Sinhala and/or English. The community owns no printed materials in the creole, though some older speakers informally adapt the English or Tamil writing system to write personal letters, etc. in creole. The group has no contact with any other creole speakers in the island. The creole in Batticaloa has for some time been losing ground to Tamil. The current picture is one of decreasing competence over successive generations: the younger a Burgher, the less likely he is to know the creole, and if he does manage a

degree of control, his speech exhibits markedly more Tamil features than that of his parents. Probably the main factor in the creole's decline is the breaking down of social barriers and the easier communication of recent years. Intermarriage, though not common, does now take place. Members of the Burgher community are in constant interaction with Tamil speakers. They live side by side, work, play, study, and worship together. Burgher children now learn Tamil at the same time they learn the creole. In the early part of this century, apparently, a child did not learn Tamil until of school age, although bilingualism has probably been common among the Burghers at least since the creole lost its status as a lingua franca.[27] Before then, of course, there would have been widespread bilingualism among the Tamils.

Nowadays even when creole speakers are talking together, conversation tends to be in Tamil if a non-creole speaking acquaintance is within earshot.[28] In the schools, especially, peer-group pressure against the use of creole is strong.

As Batticaloa has for some time been well connected to the rest of the island by rail and road, Burghers have left the area for extended periods for the sake of study or

---

[27] Cf. the 1880 quote from Coelho, p. 23 above.

[28] This is not a hard and fast rule; my chief consultant, for example, normally spoke creole when adressing any member of the Burgher community, no matter who else was present.

employment. While away, they have often lost their control of the creole, and on their return have not regained it. As well as diminishing the number of creole speakers, this tends to further reduce the number of possible situations to which the language is appropriate: now even when only Burghers are present, the language spoken is often Tamil.

No figures are available on how many of Batticaloa's Burghers speak creole. Most certainly understand it, but in many families Tamil has taken over as the home language. Sri Lanka Portuguese has proved remarkably tenacious in the past, putting the lie to many predictions of imminent demise. Even so, current prospects for its survival can only be described as bleak.

CHAPTER II: SYNCHRONIC PHONOLOGY OF BATTICALOA PORTUGUESE

In this chapter the phonological system of Batticaloa Portuguese is examined from a generative point of view. The treatment here will be far from complete: in particular, we shall not attempt to investigate much beyond the word level. The model for this description is in the main that of SPE, chapters I - III especially. In recent literature much discussion has centered on the degree of abstractness allowable in phonological descriptions. The general drift of opinion has been away from the excesses of abstraction embodied in the works of the late 1960's, such as SPE itself, Schane (1968), and Harris (1969), and towards more 'concreteness' in descriptions. I am basically in agreement with the attempts, begun by Kiparsky (1968) to limit the abstractness of phonological representations and accordingly differ from the orientation of SPE in some important respects.

There is no motivation in Batticaloa Portuguese for cycles of rules which operate below the level of the word. Readjustment rules have also been dispensed with and it has been assumed that the lexical representation of a word contains all the information needed for the proper operation of the Phonological rules. Finally, we have adhered to Kiparsky's earlier alternation condition, which has the effect of prohibiting rules of absolute neutralization. The result of these modifications has been to place on the

lexicon the burden of accounting for sporadic alternations, exceptions to rules, and other irregularities.

The first two sections of this chapter give an overview of the phonology: the inventory of underlying units is presented in Section 1; the surface (phonetic) distribution and syllable structure is outlined in Section 2. In the remainder of the chapter the phonological system is discussed in detail, and rules are given for the phonetic realization of underlying phonological representations. Section 3 deals with stress and vowel length, Section 4 with vowel quality and Section 5 with consonants.

1 Phonological Inventory[1]

Table 1 presents the consonants of Batticaloa Portuguese (BP) in traditional articulatory terms; Table 2 does the same for the vowels.

These are the abstract underlying segments which will appear in our phonological representations (enclosed in slashes: /.../). These elements are not intended to be bi-unique.[2] They do, however, all appear on the surface, and a traditional bi-unique analysis would result in the

---

[1] In keeping with the usual conventions for transcribing the languages of South Asia, as well as for typing convenience, the palatal affricates will be represented in phonemic notation as /c j/ rather than /č ǰ/. Phonetic representations will, however, use the more widely accepted notations. Thus without the application of any rule /c j/ will be realized as [č ǰ].

[2] As it happens, most of the consonants are.

Table 1. Consonants

		bi-labials	labio-dentals	dento-alveolars	post-alveolar	palatals	velars
stops	tns	p		t			k
	lax	b		d			g
affricates	tns					c	
	lax					j	
spirants	tns		f	s			
	lax			z			
nasals		m		n		ñ	
median resonant					r		
lateral liquid				l			
glides		w				y	

Table 2. Vowels

	front		central		back	
	short	long	short	long	short	long
high	i	i:			u	u:
mid	e	e:	ə		o	o:
low	æ	æ:		a:	ɔ	ɔ:

same set of underlying segments.[3] Their distribution would, of course, be different. A distinctive feature analysis of this system yields the matrix of Table 3.

We have omitted from consideration segments which occur in loans from Batticaloa Tamil (BT) or English (E) (probably via BT in many cases), even though these loans may be partially assimilated. The phonology of BT will be contrasted with that of Batticaloa Portuguese in Chapter III. Dutch (D) loans, however, are considered for the purposes of the present discussion to form part of the native vocabulary, despite the fact that they too are often incompletely assimilated. The native/non-native distinction drawn here is in line with several facts. BT is a source of continuing influence on the whole of the grammar, as is English to a lesser degree on the lexicon. The period of Dutch influence, however, is over. Every BP speaker knows BT and at least some English, but no BP speaker knows Dutch. To account for BT and English loans would require many additional phonemes and rules not needed for the description of words of Portuguese origin, but little additional phonological apparatus is needed to account for Dutch loans.[4]

---

[3] The status of [ŋ] would be debatable, as it is in the generative analysis presented here.

[4] The one additional phoneme needed, short /ɔ/, actually fills a hole in the pattern by completing the paired series of short and long vowels. Dutch loans are also responsible for the introduction of several new consonant clusters and for the extension of the distribution of short vowels, and of several consonants and clusters. Further details can be found in Section 2.

Table 3. Distinctive Feature Composition of Batticaloa Portuguese Phonemes

	i:	i	e:	e	æ:	æ	a:	ə	ɔ:	ɔ	o:	o	u:	u	y	w	r	l	p	b	m	f	t	d	n	s	z	c	j	ñ	k	g
Syllabic	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Sonorant	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	−	+	−	−	−	+	−	−	−	−	+	−	−
Consonantal	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+	+	+	+	+	+	+
High	+	+	−	−	−	−	−	−	−	−	−	−	+	+	+	+	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+	+
Back	−	−	−	−	−	−	+	+	+	+	+	+	+	+	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+
Low	−	−	−	−	+	+	+	−	+	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Anterior	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−
Coronal	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+	−	−	−	−	+	+	+	+	+	+	+	−	−	−	−
Round	−	−	−	−	−	−	−	−	+	+	+	+	+	+	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Long	+	−	+	−	+	−	+	−	+	−	+	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Voiced	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	+	+	−	−	+	+	−	+	−	+	+	−	+
Nasal	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	+	−	−	−	−	+	−	−
Lateral	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Continuant	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	−	−	+	−	−	−	+	+	−	−	−	−	−

Dutch loans form a closed set of items which are indistinguishable to the native BP speaker from those of Portuguese origin. BT and English loans, however, form an open set of items which are easily recognizable to native speakers of BP as foreign by virtue of their control of the systems from which these items are drawn. The slight changes undergone by some of these words when used in BP do not obscure their alien character.

2 Distribution

In this section we shall describe the word level surface distribution of segmental units, ignoring free variants. We shall not attempt to formulate either Morpheme Structure Conditions[5] or Systematic Phonetic Constraints. Rather, we shall simply present the distributions by means of tabular displays followed by examples. Hopefully, these data will be sufficient to enable the interested reader to work out the formal details of the associated SPC's.

2.1 Consonants

It will be useful to describe the distribution of consonants and consonant clusters in terms of syllable structure. Using the usual techniques[6] one can fairly easily locate

---

[5] Recent work seems to indicate that MSC's are, in any case, probably irrelevant to grammatical description; cf. Clayton 1976.

[6] Briefly, the basic principles are as follows:

i Any word initial consonant or cluster is by definition a syllable initial.

ii Any word final consonant or cluster is by definition a syllable final.

syllable boundaries unambiguously in BP. The results of this analysis are given in Table 4 (divided into parts a and b): possible syllable initial consonants and clusters are shown along the top, possible syllable final consonants and clusters down the left margin[7] with positionally determined phonetic variants (allophones in the traditional sense) grouped together.[8] The second and third horizontal lines indicate which syllable initials actually occur word initially or intervocalically respectively. The second column of 4a indicates which of the syllable finals actually occur word finally. The main body of the table

---

iii Single intervocalic consonants are assumed to be syllable initial rather than syllable final

iv Every intervocalic cluster is to be classified as a syllable initial or as a combination of a syllable final and a syllable initial. Much of the time this can be done unambiguously: i.e. there is a unique pair X and Y, where X is a syllable final and Y a syllable initial, such that the cluster in question is XY. In other cases a and b below come into play.

a If a given intervocalic cluster unambiguously has the structure XY, where X has been established as a syllable final, then Y may be taken to be a syllable initial even though it may not have been previously established as such.

b Similarly, if a given intervocalic cluster unambiguously has the structure XY, where Y has been established as a syllable initial, then X may be taken to be a syllable final.

When more than one analysis, or no analysis, is possible by the above means, decisions must be made on other grounds, such as patterning or universal principles. Some decisions will inevitably be ad hoc.

[7]Note that all non-initial syllables are consonant initial: word internal sequences of two vowels do not occur.

[8]As mentioned, free variants are ignored. Information about these and their conditioning factors will be found in Sections 3 - 5 of this chapter.

Table 4a. Surface Distribution of Consonants and Consonant Clusters

syllable initial		p	b	t	d	č	ǰ	k	g	f	s	z	m	n/ṇ	ñ	r/Dr	l/l̥	v/w/w̥	y
word initial		p	b	t	d	č	ǰ	k	g	f	s		m	n		r	l	v	
intervocalic		p	b	t	d	č	ǰ	k	g	f	s	z	m	n/ṇ	ñ	r	l/l̥	w/w̥	y
syllable final	word final																		
m	m	mp	mb								ms								
ŋ/n/N/ṇ/ñ/ŋ	n			nt	nd	nč	nǰ	ŋk	ŋg	ŋf	ns	nz				NDr			
r	r	rp		rt	rd	rč		rk	rg	rf	rs	rz		rn			rl	rw	
l/l̥	l/l̥	lp		lt	ld		lǰ	lk	lg	lf	ls		lm			(lr)		lw	
w/w	w/w			wt	w/wd	wč		wk	wg			wz		wn		w/wr			
y	y			yt											yñ	yr	yl	yv	yy
s	s	sp		st				sk											
z									zg				zm						
p																			
b	b																		
t	t															*			
k	k			kt												*			
f															fn				
yn	yn																		
yl					yl+d														
yt	yt																		
wr				wrt							wrs								
ws	ws																		
nt	nt																		

Notes:

* indicates that the cluster in question can be found in Table 4b.

\+ indicates that the cluster in question occurs only across a morpheme boundary

Table 4b. Surface Distribution of Consonants and Consonant Clusters (cont.)

syllable initial	sp	st	sk	pr	br	tr	dr	kr	gr	fr	wr	str	skr	pl	bl	kl	gl	fl	sl	kw	gw	sw	drw	frw	fs
word initial	sp	st	sk	pr	br	tr	dr	kr	gr	fr		str		pl	bl	kl		fl	sl	kw		sw		frw	
inter-vocalic	*	*	*		br	tr	dr	kr	gr	fr	*	*	*										d+rw		
syllable final																									
m				mpr	mbr									mpl	mbl										
ŋ/n/ŋ	nsp		nsk			ntr	ndr		ŋgr				nskr				ŋgl			ŋk+w	ŋgw				ŋfs
l											lwr														
y			ysk																						
s				spr		str		skr																	
p		pst																							

Notes:

* indicates that the cluster in question can be found in Table 4a.

+ indicates that the cluster in question occurs only across a morpheme boundary

presents the occurring intervocalic clusters analyzed as combinations of possible syllable finals and syllable initials.[9] The class of ambiguous intervocalic clusters represented by the schema sC(r) has been resolved as s$C(r) (where $ indicates the syllable division) since word initially such clusters are rare, especially in words derived from Portuguese. Similarly intervocalic tr and kr have been analyzed as $tr and $kr respectively, since t and k are otherwise extremely rare as syllable finals.

Examples:

##pV	(1)	[pé:si]	'fish'
##bV	(2)	[bé:su]	'lip'
##tV	(3)	[tɔ́:si]	'cough' N
##dV	(4)	[dɔ́:si]	'sweet'
##čV	(5)	[čú:wə]	'rain'
##ǰV	(6)	[ǰuwá•]	'gamble' V
##kV	(7)	[kəstigá•]	'bother' V
##gV	(8)	[gəstá•]	'spend; waste'
##fV	(9)	[fí:yu]	'son'
##sV	(10)	[suwɔ́•r]	'sweat'
##mV	(11)	[má:s]	'more'
##nV	(12)	[ná:]	'no, (NEG)'

[9] It goes without saying that the reliability of this entire work depends on the accuracy and representativity of the corpus. This limitation is especially relevant to the distributional statements in this section. The size of my corpus (approximately 6000 entries) ensures a good degree of representativity while increasing the possibility of human error.

##rV		(13)	[rǽ:ñə]	'snot'
##lV		(14)	[lǽ:ñə]	'firewood'
##vV		(15)	[ví:ñu]	'arrack'
##spV	one ex. only	(16)	[spəñó·l]	'Spanish'
##stV	one ex. only	(17)	[stém]	'trunk, stem' < D
##skV	one ex. only[10]	(18)	[skərní·l]	'hinge' < ?D
##prV		(19)	[prǽ:ndu]	'education'
##brV		(20)	[brí:ŋku]	'game'
##trV		(21)	[trí:pə]	'intestines'
##drV		(22)	[dré:tu]	'right, correct'
##krV		(23)	[kriyá·nsə]	'child'
##grV		(24)	[grá:ǰə]	'crow'
##frV		(25)	[frí:w]	'cold'
##strV	one ex. only[11]	(26)	[stroná·y]	'bridesmaid' < ?D
##plV	one ex. only	(27)	[plét]	'flat' < D
##blV	D loans only	(28)	[blét]	'tin'
##klV	rare	(29)	[klá:r]	'clear'
##flV	rare	(30)	[flǽ:mə]	'phlegm'
##slV	D loans only	(31)	[slǽktə]	'bad'
##kwV	rare	(32)	[kwæ:ntru]	'coriander'
##swV	one ex. only	(33)	[swá:r]	'excessive' ?
##frwV	one ex. only	(34)	[frwí:tə]	'fruit'
V$pV		(35)	[kupá·]	'stare; look after'
V$bV		(36)	[kəbé·lu]	'hair'

[10] Also in [ski], a post-clitic marking reported speech.

[11] Assuming [stronǰá·ŋgəl] 'best man' contains another instance of the same formative.

V$tV		(37)	[pé:tu]	'chest'
V$dV		(38)	[dé:du]	'finger'
V$čV		(39)	[fé:či]	'lock' N
V$ǰV	rare	(40)	[gré:ǰə]	'church'
V$kV		(41)	[sǽ:ku]	'dry' A
V$gV		(42)	[sǽ:gu]	'blind'
V$fV		(43)	[bɔ́:fi]	'lungs'
V$sV		(44)	[dɔ́:si]	'sweet'
V$zV		(45)	[dɔ́:zi]	'twelve'
V$mV		(46)	[iská•mu]	'(fish) scale' Cf. SP escama
V$nV		(47)	[kurtí•nə]	'mosquito net'
V$n̤V		(48)	[tɔ́:n̤ə]	'afterwards'
V$ñV		(49)	[kuzí•ñə]	'kitchen'
V$rV		(50)	[ərá•]	'miss; go wrong'
V$lV		(51)	[pilá•]	'pound (rice, pulse, etc.)'
V$l̤V		(52)	[əl̤á•]	'there'
V$wV		(53)	[əwá•nə]	'fan'
V$wV		(54)	[kuwá•]	'strain'
V$yV		(55)	[əyá•]	'clean (rice, pulse, etc.)'
V$spV	see Vs$pV			
V$stV	see Vs$tV			
V$skV	see Vs$kV			
V$brV	rare	(56)	[obrigá•du]	'thanks'
V$trV		(57)	[ká:tru]	'four'

V$drV		(58)	[pǽ:drə]	'stone'
V$krV		(59)	[sú:kri]	'sugar'
V$grV		(60)	[tí:gri]	'tiger'
V$frV		(61)	[bú:frə]	'(water) buffalo'
V$vrV	see Vv$rV			
V$strV	see Vs$trV			
V$skrV	see Vs$krV			
V$drwV	one ex. only	(62)	[lədrwí•sə]	'dishonesty' (=/lədr+w+i:s+ə/)
Vm##		(63)	[pá:m]	'bread'
Vm$pV		(64)	[lí:mpu]	'clean'
Vm$bV		(65)	[pó:mbə]	'dove'
Vm$sV	one ex. only	(66)	[kumsá• ~ kunsá•]	'begin'
Vm$prV		(67)	[kumprí•du]	'long'
Vm$brV		(68)	[kəmbrá•m]	'prawn'
Vm$plV	one ex. only	(69)	[komplumá•s]	'maldive fish' < ? Cf. BT /ma:ci/ ([má:ši]) id.
Vm$blV	one ex. only	(70)	[nǽ:mbli]	'ridged bowl for cleaning rice, pulse, etc.'
Vn##		(71)	[iskruwá•n]	'writer'
Vn$tV		(72)	[ɔ́:ntə]	'yesterday'
Vn$dV		(73)	[ɔ́:ndə]	'wave' N
Vñ$čV		(74)	[íñčə]	'draw (water)'

V$\underline{\tilde{n}}$\$\underline{\check{j}}$V		(75)	[uñǰuwé˙y]	'knee'[12]
V$\underline{\eta}$\$\underline{k}$V		(76)	[brí:ŋku]	'game'
V$\underline{\eta}$\$\underline{g}$V		(77)	[lí:ŋgu]	'tongue; language'
V$\underline{\eta}$\$\underline{f}$V		(78)	[koŋfiyá˙]	'hope' V
V$\underline{n}$\$\underline{s}$V		(79)	[lǽ:nsu]	'handkerchief'
V$\underline{n}$\$\underline{z}$V		(80)	[ɔ́:nzi]	'eleven'
V$\underline{N}$\$\underline{Dr}$V	rare	(81)	[tǽ:NDru]	'unripe coconut'
V$\underline{n}$\$\underline{sp}$V	one ex. only	(82)	[sənspí˙rə]	'a type of fish' <? (=BT/kiLakkan/)
V$\underline{n}$\$\underline{sk}$V	rare	(83)	[á:nski]	'before' Cf. SP antes que
V$\underline{n}$\$\underline{tr}$V		(84)	[dǽ:ntru]	'inside'
V$\underline{n}$\$\underline{dr}$V	one ex. only	(85)	[pindrá˙]	'hang'
V$\underline{\eta}$\$\underline{gr}$V		(86)	[kəŋgré˙y]	'crab'
V$\underline{n}$\$\underline{skr}$V	one ex. only	(87)	[bənskrú˙]	'vise' <D?

---

[12]Although this is primarily a synchronic study, the historical derivation of this form deserves mention, as it is the regular development of SP um joelho ([ũ ǰuéĨu]) 'a/one knee.' Cf.:

BP [duwé˙nti] 'ill' SP doente ([duéti]) id.

BP [vǽ:y] 'old' SP velho ([véĨu]) id.

BP [uñ ǰudé˙w] 'a/one Jew' SP um judeu ([ũ ǰudéu]) id.

Fusion of the article in (75) probably resulted from the frequent occurrence of the form in the expression

BP / $[[\text{unjuve:y}]_{N}\text{ntu}]_{NP}[\text{sənt+a:}]_{V}$/

(knee LOC sit) 'sit on (your) knee; i.e. kneel'

The original intent of the expression, viz. 'sit on one knee' is clear.

V$\underline{ŋ}$\$$\underline{gl}$V	rare	(88)	[iŋglé•s ~ iŋgré•s]	'English'
V$\underline{ŋ}$\$$\underline{kw}$V	one ex. only	(89)	[siŋkwǽ•ntə]	'fifty' (=/si:nk+w+æ:nt+ə/
V$\underline{ŋ}$\$$\underline{gw}$V	rare	(90)	[məŋgwá•du]	'faded' <?
V$\underline{ɱ}$\$$\underline{fs}$V	one ex. only	(91)	[koɱfsá•]	'confess'
V$\underline{r}$##		(92)	[suwó•r]	'sweat' N
V$\underline{r}$\$$\underline{p}$V		(93)	[kó:rpu]	'body'
V$\underline{r}$\$$\underline{t}$V		(94)	[kortá•]	'cut, slaughter'
V$\underline{r}$\$$\underline{d}$V		(95)	[kó:rdə]	'rope'
V$\underline{r}$\$$\underline{č}$V	one ex. only	(96)	[pófərči]	'poffertje: a fritter-like sweetmeat' <D
V$\underline{r}$\$$\underline{k}$V		(97)	[pó:rku]	'pig'
V$\underline{r}$\$$\underline{g}$V		(98)	[lá:rgu]	'wide'
V$\underline{r}$\$$\underline{f}$V		(99)	[gá:rfu]	'fork'
V$\underline{r}$\$$\underline{s}$V		(100)	[pórsə]	'see, appear'
V$\underline{r}$\$$\underline{z}$V		(101)	[kətó•rzi]	'fourteen'
V$\underline{r}$\$$\underline{n}$V		(102)	[kó:rnu]	'horn'
V$\underline{r}$\$$\underline{l}$V	one ex. only	(103)	[orló•zu]	'clock'
V$\underline{r}$\$$\underline{w}$V		(104)	[bá:rwə]	'beard'
V$\underline{l}$##		(105)	[mǽ:l]	'honey'
V$\underline{l̥}$##		(106)	[má:l̥]	'bad'
V$\underline{l}$\$$\underline{p}$V		(107)	[kú:lpə]	'guilt'
V$\underline{l}$\$$\underline{t}$V		(108)	[á:ltu]	'tall, high'
V$\underline{l}$\$$\underline{d}$V	rare	(109)	[ká:ldu]	'gravy, sauce'
V$\underline{l}$\$$\underline{ǰ}$V	one ex. only	(110)	[əlǰó•fri]	'pearl'
V$\underline{l}$\$$\underline{k}$V	rare	(111)	[əlkonsá•]	'consult' <?

VlgV	rare	(112)	[mǽ:lgə]	'mosquito'
VlfV	rare	(113)	[əlfá·də]	'pillow'
VlsV	one ex. only	(114)	[kəlsá·m]	'trousers'
VlmV	rare	(115)	[ælmǽ·ru]	'wardrobe, cupboard'
Vl(w)rV	one ex. only	(116)	[á:lwri ~ á:lri]	'tree, plant'
VlwV	rare	(117)	[pilwá·]	'beat'
Vw##		(118)	[pá:w]	'plank'
Vw##		(119)	[ó:w]	'egg'
VwtV	rare	(120)	[bəwtí·zmu]	'baptism'
VwdV		(121)	[á:wdu]	'sharp'
VwdV	one ex. only	(122)	[kó:wdu]	'cubit'
Vwč	one ex. only	(123)	[réwči]	'rib'
VwkV	one ex. only	(124)	[rəwkí·n]	'violin'
VwgV	one ex. only	(125)	[əwgá·]	'fly' V
VwzV	one ex. only	(126)	[ləwzí·n]	'king coconut' <?
VwnV	one ex. only	(127)	[əwná·]	'fan' V
VwrV		(128)	[kɜwrá·]	'break'
VwrV	rare	(129)	[kuwrí·]	'cover' V
Vy##		(130)	[pá:y]	'father'
VytV	rare	(131)	[ó:ytu]	'eight'
VyñV	rare	(132)	[məyñá·tu]	'washerman'
VyrV		(133)	[rozá·yru]	'rosary'
VylV	one ex. only	(134)	[bá:ylu]	'dance' N
VywV	one ex. only	(135)	[rá:ywə]	'anger'
VyyV	one ex. only	(136)	[búyyu]	'monkey'

Vy$skV	one ex. only	(137)	[məyská·rdə]	'left'[13]
Vs##		(138)	[lé:s]	'read' V
Vs$pV		(139)	[ispí·m]	'fine bone; quill; thorn'
Vs$tV		(140)	[tǽ:stə]	'head'
Vs$kV		(141)	[muskí·tu]	'fly' N
Vs$prV	one ex. only	(142)	[isprú·mə]	'squeeze, wring'
Vs$trV		(143)	[istrǽ·lə]	'star'
Vs$krV		(144)	[iskrú·wə]	'write'
Vz$gV	rare	(145)	[rezgətá·]	'save'
Vz$mV		(146)	[frí:zmu]	'dear' < SP felíssimo ?
Vp$stV	one ex. only	(147)	[kəpstá·y]	'ability'
Vb##	one ex. only	(148)	[nɔ́b]	'knot' < D?
Vt##	D loans only	(149)	[blɐ́t]	'tin'
Vk##	one ex. only	(150)	[dǽk]	'storey' < D
Vk$tV	one ex. only	(151)	[slǽktə]	'bad' < D
Vf$nV	one ex. only	(152)	[ǰá:fnə]	'Jaffna (a place name)' (=BT/ya:Lpa:Nam/ S/ya:pane:/)
Vyn##	?D loans only	(153)	[déyn]	'inch'
Vyl$dV	one ex. only	(154)	[bəyldó·r]	'dancer' (=/ba:yl+do:r/)

---

[13] Cf. SP mão esquerda 'left hand'
BP /[məyska:rd+ə]$_A$[ma:m]$_N$/ id.

V<u>wr</u>$<u>t</u>V	one ex. only (155)	[gəwrtá•]	'dig'
V<u>wr</u>$<u>s</u>V	one ex. only (156)	[əwrsá•]	'hug'
V<u>ws</u>##	(157)	[dé:ws]	'god'
V<u>nt</u>##	one ex. only (158)	[rónt]	'round' < D

Not classified in Table 4 are clusters which can occur across word boundaries, including clusters resulting from the transformational addition of inflectional affixes: for example any of the word final clusters (except <u>nt</u> and <u>yt</u>, which do not occur in nouns, and <u>m</u>, which assimilates to [n]) may occur before the genitive suffix #<u>su</u>:

(159) /kəklu:n#su/ ⟶ [kəlkú•nsu] 'turkey's'
(160) /a:s#su/ ⟶ [á:ssu] 'steel's'
(161) /dəyn#su/ ⟶ [dáynsu] 'inch's'

Similarly we omit from consideration intervocalic [ŋ] derived from /ng#(#)V/. See Section 5 for further details.

2.2 Vowels

BP has retained the seven oral vowels of SP but has lost the distinction between nasalized and non-nasalized vowels. A word may contain at most one phonetically long vowel, which is always stressed; such vowels are longer in initial than in non-initial syllables. Short stressed vowels also occur in initial syllables, but not in non-initial syllables. Unstressed short vowels are found in all syllables. The distribution of the vowels is summarized in Tables 5a and 5b. A detailed analysis of stress and vowel length will be

Table 5a. Distribution of Long (and Stressed) Vowels

	i:/ɨ:/ i•/ɨ•	e:/ɜ:/ e•/ɜ•	æ:/æ•	a:/a•	ɔ:/ɔ•	o:/o•	u:/u•
I initial syllable vowels							
A in monosyolables							
i) _ C_1	i:	e:	æ:	a:	ɔ:	o:	u:
ii) _ #	i:	e:		a:	ɔ:	(o:)	u:
B in disyllables	i:/ɨ:	e:/ɜ:	æ:	a:	ɔ:	o:	u:
C in polysyllables	i:		æ:	a:	ɔ:		u:
II medial syllable vowels	i•	e•	æ•	a•	ɔ•	o•	u•
III final syllable vowels							
i) _ C_1	i•	e•	æ•	a•	ɔ•	o•	u•
ii) _ #	i•	e•		a•		o•	u•

Table 5b. Distribution of Short Vowels

	i/ɨ	e/ɜ*	æ	ə	ɔ	o	u
I initial syllable vowels							
A in monosyllables (stressed; _C_1)		e	æ	ə	ɔ	o	**
B in disyllables i) stressed	i/ɨ	e/ɜ	æ	ə	ɔ	o	u
ii) unstressed	i/ɨ	e/ɜ	æ	ə		o	u
C in polysyllables i) stressed	i	e	æ	ə			
ii) unstressed	i/ɨ	e/ɜ		ə		o	u
II medial syllable vowels (unstressed)	i	e/ɜ		ə		o	u
III final syllable vowels (unstressed) i) _ C_1	i			ə			
ii) _ #	i			ə			u

* Though [ɜ] is a free variant of [e] wherever the latter occurs, only those positions in which obligatory [ɜ] before rC is found are indicated. See Section 4.3 for a more detailed statement.

**Unstressed [u] occurs in the post-clitic [tu] '(perfective)'.

found in Section 3.

Examples:

a. Long (Stressed) Vowels

I initial syllable vowels

A in monosyllables

i) _C_1

i:	rare	(162)	[fí:w]	'plaintain, banana'
		(163)	[mí:l]	'thousand'
e:	rare	(164)	[fé:w]	'ugly'
		(165)	[ké:m]	'who'
æ:	rare	(166)	[prǽ:w]	'nail' N
		(167)	[mǽ:l]	'honey'
a:		(168)	[á:w]	'water'
		(169)	[dá:y]	'hit'
		(170)	[pá:m]	'bread'
ɔ:	rare	(171)	[vɔ́:s]	'voice'
		(172)	[sɔ́:m]	'sound, noise'
o:	rare	(173)	[ó:y]	'eye'
		(174)	[nó:s]	'we'
u:	one ex. only	(175)	[sú:y]	'dirt, dust'

ii) _#

i:	rare	(176)	[fí:]	'rod'
		(177)	[ví:]	'come'
e:	rare	(178)	[ré:]	'king'
		(179)	[pé:]	'foot'
a:	one ex. only	(180)	[dá:]	'give'

ɔ: rare	(181) [sɔ́:]	'alone'
	(182) [pɔ́:]	'powder'
(o: only in the postclitic	(183) [vó:]	'(indefinite)')
u: rare	(184) [kú:]	'ass'
	(185) [bú: ~ bú:wa]	'elder brother'

B in disyllables

i:	(186) [fí:yu]	'son'
	(187) [pí:ris]	'saucer'
ɨ:rC rare	(188) [ɨ́:rti]	'straight'
	(189) [vɨ́:rzim]	'(The) Virgin'
e:	(190) [mé:yu]	'half'
	(191) [gré:ǰə]	'church'
ʒ:rC one ex. only	(192) [vʒ́:rdi]	'green'
æ:	(193) [vǽ:yə]	'old woman'
	(194) [lǽ:ñə]	'firewood'
	(195) [mǽ:nis]	'less'
a:	(196) [grá:ǰə]	'crow' N
	(197) [lá:ñə]	'young coconut'
	(198) [vá:rzim]	'cultivation; field'
ɔ:	(199) [fɔ́:yə]	'leaf'
	(200) [pɔ́:rtə]	'door'
	(201) [mɔ́:li]	'soft'
o:	(202) [čó:ru]	'crying'
	(203) [mó:rtu]	'dead'

		(204) [mó:vis]	'clouds'[14]	
u:		(205) [čú:wə]	'rain'	N
		(206) [kú:rtu]	'short'	

C in polysyllables

i:	one ex. only	(207) [ǰi:ñǰivə]	'gums'	
æ:	one ex. only	(208) [ǽ:tikə]	'tuberculosis'	
a:	rare	(209) [á:n̥imə]	'courage'	
		(210) [á:viyəs]	'birds'	
ɔ:	rare	(211) [gɔ́:mitu]	'vomit'	N
		(212) [glɔ́:riyə]	'glorious; heaven'	
u:	one ex. only	(213) [lú:wriyə ~ lú:griyə]	'rent'	N

II medial syllables

i˙	(214) [sənspí˙rə]	'a type of fish' =BT /kiLakkan/	
	(215) [muskí˙tu]	'fly'	N
e˙	(216) [esté˙rə]	'mat'	
	(217) [əzé˙ti]	'oil'	
æ˙	(218) [oytǽ˙ntə]	'eighty'	
	(219) [kærǽ˙tə]	'cart'	
a˙	(220) [səpá˙tu]	'sandal'	
	(221) [lumá˙rə]	'moonlight'	

---

[14] (204) is an exception to the general distribution of [ɔ́:] and [ó:] in initial syllables:

[ɔ́:] / _ C_1 $\{^{ə}_{i}\}$ [ó:] / _ C_1 u

ɔ·	(222) [siñɔ́·rə]	'(Our) Lady'	
	(223) [impɔ́·rtə]	'importance'	
o·	(224) [səbó·di]	'health'	
	(225) [koró·və]	'crown' N	
u·	(226) [vɜrdú·rə]	'vegetable'	
	(227) [kərdəmú·ŋgu]	'cardamom'	

III final syllable vowels

i) _C_1

i·	(228) [nərí·s]	'nose'	
	(229) [pikiní·m]	'small'	
e· rare	(230) [ǰudé·w]	'Jew'	
	(231) [kəŋgré·y]	'crab'	
æ· rare	(232) [mursǽ·w]	'bat'	
	(243) [mæsǽ·w]	'young man'	
a·	(234) [kintá·l̥]	'garden; compound'	
	(235) [əruməsá·m]	'roof'	
ɔ· rare	(236) [ənənɔ́·s]	'pineapple'	
	(237) [spəñɔ́·l̥]	'Spanish'	
o·	(238) [dispó·s]	'afterwards'	
	(239) [kəló·r]	'heat'	
u· rare	(240) [əzú·l]	'blue'	
	(241) [səl̥ú·s]	'hiccup'	

ii) _#

i·	(242) [repɜrtí·]	'share' V	
	(243) [əwrí·]	'open' V	

e•[15]	one ex. only	(244)	[mɜrsé•]	'thanks'
a•		(245)	[əlá•]	'there'
		(246)	[eŋkontrá•]	'meet'
o•	one ex. only	(247)	[kəlkó•]	'a type of fish' = BT /koTuva/
u•	one ex. only	(248)	[bəmbú•]	'bamboo'

b. Short Vowels[16]

I initial syllable vowels

A in monosyllables (stressed; $\_C_1\#$)

e	one ex. only	(249)	[krél]	'curl' < D
æ	one ex. only	(250)	[dǽk]	'storey' < D
ə	rare	(251)	[kə́y]	'fall' V
		(252)	[blə́t]	'tin' < D
ɔ	D loans only	(253)	[rɔ́nt]	'round'
		(254)	[nɔ́b]	'knot'
o	one ex. only	(255)	[óy]	'today'
(u	see note to Table 5b p. 54.)			

---

[15] Before a word boundary, /ə/ has the free variant [ę-] (/e/ does not occur in this position). When stressed because of discourse considerations (contrast, emphasis, continuity, rhythm of a song, etc.), it becomes [é•]. E.g. I have observed a speaker say [kuziñé•] for /kuzi:ñ+ə/ 'kitchen' (normally [kuzí•ñə]) in order to contrast it with /kuzi:ñ+a:/ 'cook' V ([kuziñá•]).

[16] The low unstressed (short) vowels contrast only minimally with their mid counterparts.
Unstressed [æ ɔ] are found only i nthe environment:

$$\#C_o\text{-}C_o \begin{bmatrix} V \\ +\text{low} \\ \alpha\text{round} \\ \alpha\text{back} \end{bmatrix}$$

Mid [e o] are relatively rare in this position. For further details, see discussion of (603) Raising and (612) Vowel Harmony in Section 4.

B in disyllables

i) stressed

i	(256) [ísti]	'this' DEM
	(257) [íñčə]	'draw (water)'
ɨrC one ex. only	(258) [dɨ́rtə]	'melt'
e	(259) [éli]	'he'
	(260) [préndə]	'learn'
ɜrC one ex. only	(261) [pɜ́rdə]	'lose'
æ rare	(262) [ǽlə]	'she'
	(263) [slǽktə]	'bad' < D
ə	(264) [čə́pə]	'hat'
	(265) [fə́yə]	'do'
ɔ one ex. only	(266) [ɔ́ŋkəl̩]	'uncle' < ?
o	(267) [bósə]	'your, yours'
	(268) [tósə]	'cough' V
u	(269) [búyyu]	'monkey'
	(270) [kúnsə]	'recognize'

ii) unstressed

i	(271) [kintá•l̩]	'compound, garden'
ɨrC one ex. only	(272) [sɨrví•]	'work' V
e	(273) [peská•]	'fish' V
ɜ rare	(274) [vɜrdá•y]	'truth'
æ rare	(275) [mæsǽ•w]	'young man'
ə	(276) [kəntá•]	'sing'
o	(277) [kontá•]	'count'
u	(278) [kunsá• ~ kumsá•]	'begin'

C in polysyllables

i) stressed

i	rare	(279)	[síndəfərə]	'Monday'
e	one ex. only	(280)	[séstəfərə]	'Friday'
æ	rare	(281)	[tærsəfərə]	'Tuesday'
ə	rare	(282)	[évərə]	'now'
		(283)	[kéndiyə]	'oil lamp'

ii) unstressed

i		(284)	[pikiní·m]	'small'
		(285)	[irigí·]	'get up'
ɨrC	one ex. only	(286)	[sɨrvidó·r]	'servant'
e		(287)	[eŋkontrá·]	'meet'
		(288)	[sesé·ntə]	'sixty'
ʒ	rare	(289)	[vʒrdú·rə]	'vegetable'
		(290)	[pʒrduwá·]	'pardon'
æ		(291)	[ælmé·ru]	'wardrobe'
		(292)	[værgó·ñə]	'shame, shyness'
		(293)	[væstumé·ntu]	'windy weather'
ə		(294)	[ənənó·s]	'pineapple'
		(295)	[kəmbrá·du]	'friend'
ɔ		(296)	[sɔvó·lə]	'onion'
		(297)	[nɔvé·ntə]	'ninety'
o		(298)	[rozá·yru]	'rosary'
		(299)	[oyté·ntə]	'eighty'
u		(300)	[luná·yru]	'lightning'
		(301)	[kumiyá·m]	'Communion'

II medial syllables (unstressed)

i	(302)	[ənimá•l]	'animal'
e	(303)	[penerá•]	'sift'
ɜrC rare	(304)	[repɜrtí•]	'share' V
ə	(305)	[əviyəmǽ•ntu]	'tool'
o	(306)	[eŋkontrá•]	'meet'
u	(307)	[iskurdá•m]	'darkness'

III final syllables (unstressed)

i) _C_1

i	(308)	[mǽ•nis]	'less'
	(309)	[vá:rzim]	'cultivation; field'
ə	(310)	[górgəl]	'throat'
	(311)	[kánəy]	'fellow (pej.)'

ii) _#

i	(312)	[má:či]	'male'
	(313)	[pé:si]	'fish' N
ə	(314)	[fí:yə]	'daughter'
	(315)	[má:čə]	'march, walk' N
u	(316)	[fí:yu]	'son'
	(317)	[bé:su]	'lip'

3 Stress Placement and Vowel Length

One of the most striking differences between Batticaloa Portuguese and Standard Portuguese is a typological difference in the status of stress and vowel length. In both languages there is a canonical (i.e. most common) stress

position: the last syllable if this is closed, otherwise the penult. In Standard Portuguese stress is not completely predictable; but given stress, vowel length is predictable, stressed vowels being slightly longer than unstressed vowels. We would therefore posit a low-level rule of vowel lengthening such as (318).

(318) Vowel Lengthening in Standard Portuguese:

$$\begin{bmatrix} V \\ +\text{stress} \end{bmatrix} \longrightarrow [+\text{long}]$$

In Batticaloa Portuguese the reverse situation obtains: as a result of various phonological changes and borrowings, long and short vowels now stand in contrast. Stress on the other hand is fully predictable. This section will be concerned primarily with the word level analysis of stress and vowel length in BP, though towards the end of the section, we shall briefly look at what happens in syntactic constructions.

Stress and vowel length are still closely interrelated in Batticaloa Portuguese: there is no restriction on the number of underlying long vowels in a word, but a word may contain at most one phonetically long vowel. All unstressed vowels are short, or to put it another way, <u>all surface long vowels are stressed. It is not true, however, that all stressed vowels are long, and in initial syllables both stressed long and stressed short vowels are found</u>. The contrasts in (319) - (332) clearly demonstrate that vowel

length is not simply a surface phenomenon and must be lexically represented. All of these items are of Standard Portuguese (SP) or Dutch (D) origin; none are from indigenous sources.

(319) a)	[dísə]	'descend'	SP descer	id.
b)	[mí:sə]	'mass'	SP missa	id.
(320) a)	[ísti]	'this (one)'	SP este	id.
b)	[trí:sti]	'sad'	SP triste	id.
(321) a)	[kúmə]	'eat'	SP comer	id.
b)	[lú:mi]	'light'	SP lume	id.
(322) a)	[úŋə]	'one'	SP um/uma	id.
b)	[ú:ñə]	'fingernail'	SP unha	id.
(323) a)	[kérə]	'want'	SP querer	id.
b)	[sé:rə]	'wax'	SP cera	id.
(324) a)	[krél]	'curl'	D krul	id.
b)	[ke:m]	'who'	SP quem	id.
(325) a)	[óy]	'today'	SP hoje	id.
b)	[ó:y]	'eye'	SP olho	id.
(326) a)	[bósə]	'your(s)'	SP vosso/a	id.
b)	[pó:su]	'well'	SP poço	id.
(327) a)	[slǽktə]	'bad'	D slecht	id.
b)	[tǽ:stə]	'head'	SP testa	id.
(328) a)	[ǽlə]	'she'	SP ela	id.
b)	[vǽ:lə]	'sail'	SP vela	id.
(329) a)	[ról̩]	'roll'	D rollen	id.
b)	[sɔ́:l̩]	'sun'	SP sol	id.

(330) a)	[gɔ́rgəl]	'throat'	D	gorgel	id.
b)	[fɔ́:rsə]	'strength'	SP	força	id.
(331) a)	[ékə]	'that (one)'	SP	aquele/a	id.
b)	[bá:kə]	'cow'	SP	vaca	id.
(332) a)	[stém]	'stem, trunk'	D	stam	id.
b)	[ča:m]	'ground, earth'	SP	chão	id.

All fourteen vowels of Batticaloa Portuguese are represented in the above list. Only for the pair [a:] : [ə], illustrated in (331) and (332), is the difference in quantity accompanied by a difference in quality: in (319) - (330) length is serving as the sole phonetic feature distinguishing two vowels.

It is only in initial syllables that length serves as a minimal phonetic feature of contrast. In non-initial syllables the situation is slightly different; here stress and length are concomitant features. That is, we find only two types of vowels: (i) unstressed short vowels and (ii) stressed vowels which vary in length from half-long to short (and which we shall call 'potentially half-long'). Examples (333) - (335) illustrate these two possibilities.

(333) a)	[kədí·yə ~ kədíyə]	'chain'	SP	cadeia	id.
b)	[kéndiyə]	'oil lamp'	SP	candeia	id.
(334) a)	[nərí·s ~ nəris]	'nose'	SP	nariz	id.
b)	[pí:ris]	'saucer'	SP	pires	id.
(335) a)	[iñčá· ~ iñčá]	'swell'	SP	inchar	id.
b)	[íñčə]	'fill'	SP	encher	id.

The length of the potentially half-long vowels is not completely freely determined: they tend, for example, to be short in closed or final syllables and in rapid, less careful speech. As length is in any case significant for initial syllable vowels, we analyze these stressed potenfially half-long vowels in non-initial syllables as underlying long vowels. This treatment is supported by the length alternations in (336) - (341). In these examples a stressed vowel in the same morpheme is either long or potentially half-long, depending on whether it is in an initial or non-initial syllable.

(336) a)	[ǒ:ytu]	'eight'
b)	[dizǒ˙ytu]	'eighteen'
(337) a)	[dí:yə]	'day'
b)	[medí:yə]	'midday'
(338) a)	[kru:wə]	'raw'
b)	[əskrú˙wə ~ əskrúwə]	'raw rice'[17] (<SP <u>arroz</u> + SP <u>crua</u>)
(339) a)	[kǒ:ntə]	'amount'
b)	[diskǒ˙ntə ~ diskǒntə]	'discount'
(340) a)	[ǒ:rə]	'hour'
b)	[əkəlǒ˙rə ~ əkəlǒrə]	'then' (<SP <u>aquela hora</u>)

[17] There may no longer be any synchronic connection between the two items in (338), though from a historical perspective they do certainly have a morpheme in common. The same may be said of (339) and (340).

(341) a) [té:m] 'is'

b) [loté·m ~ lotém] 'will be'

c) [nunté·m ~ nuntém] 'is not'

In each of these, then, our analysis would posit underlying long vowels which become stressed (the stress placement rule will be given below), and a later scalar rule would then adjust the length of stressed vowels in non-initial syllables. Assuming that long vowels are phonetically [2long] and half-long vowels are [1long] and that [0long] = [-long], such a scalar rule might look like (342).

(342) Length Adjustment (LA): first approximation

$$\left[\begin{smallmatrix} V \\ +\text{stress} \end{smallmatrix}\right] \longrightarrow [n\text{long}] \;/\; C_o V C_1 \_$$

where n is 1 or 0 and is a function of phonetic environment, speech style, etc.

Having established the phonemic status of vowel length, we are now in a position to predict stress placement: stress falls on the last underlying long vowel of a word, or on the first vowel of a word having no long vowel. A rule such as (343) will accomplish this.

(343) Main Stress Rule (MSR): first approximation

$$V \longrightarrow [+\text{stress}]$$

$$/\ [\ \langle X \rangle\ C_o\ \left[\begin{smallmatrix} \overline{\quad} \\ \langle +\text{long} \rangle \end{smallmatrix}\right]\ (C_o\left[\begin{smallmatrix} V \\ -\text{long} \end{smallmatrix}\right]C_o)_o\ ]_L$$

where L ranges over the lexical categories.[18]

[18] I.e. L=N,V,etc. Certain clitics and particles are never stressed and presumably must be excluded from the

The rule is an abbreviation of two disjunctively ordered rules: the first rule tries to stress a long vowel which is followed by no other long vowels in the word (i.e. it will stress the last long vowel in the word); if there is no long vowel in the word it fails to apply, and the second rule places stress on the first vowel of the word.

We first illustrate stress placement in the simplest case: that of a word having at most one underlying long vowel. There are three classes of such words:

i Words with no long vowels: the initial syllable vowel is stressed and short. This class contains most disyllabic verbs ending in ə (historically from SP verbs in -er), most disyllabic pronouns and determiners, a few other historically Portuguese words, and many items of Dutch or uncertain origin.

(344) /oy/ ——> [óy] 'today'

(345) /kəy/ ——> [kéy] 'fall' V

(346) /rəy/ ——> [réy] 'line' N

(347) /krel/ ——> [krél] 'curl (N), curly'

(348) /stəm/ ——> [stém] 'stem, trunk'

(349) /blət/ ——> [blét] 'tin'

---

range of L:
/kə(n)di/ '(intensifier)' /fəl+a:(#tu)/ '(quotative)'
/ley/ 'like' /ski/ '(attributed speech)'
/na:/ '(tag question)' /tu/ '(perfective)'
Others, like /ta:m/ 'also', do get stressed.
Still others, like /se:/ '(conditional)', are optionally stressed.

(350) /rɔnt/ ——> [rɔ́nt] 'round'

(351) /dæk/ ——> [dǽk] 'storey'

(352) /kəws/ ——> [kə́ws] 'sp. of fish' (= BT /uLuvay/)

(353) /ək+ə/ ——> [ə́kə] 'that, that one'

(354) /ist+i/ ——> [ísti] 'this, this one'

(355) /miñ+a/ ——> [míñə] 'my'

(356) /el+i/ ——> [éli] 'he'

(357) /æl+ə/ ——> [ǽlə] 'she'

(358) /botus/ ——> [bótus] 'you [HON/PL]'

(359) /etus/ ——> [étus] 'they [HON]'

(360) /kuns+ə/ ——> [kúnsə] 'recognize, make out'

(361) /bew+ə/ ——> [bɜ́və] 'drink'

(362) /trim+ə/ ——> [trímə] 'tremble'

(363) /dis+ə/ ——> [dísə] 'descend, get down'

(364) /cəp+ə/ ——> [čə́pə] 'hat'

(365) /pəw+i/ ——> [pə́vi] 'wick'

(366) /buyyu/ ——> [búyyu] 'monkey'

(367) /slækt+ə/ ——> [slǽktə] 'bad'

(368) /slengər/ ——> [slé ŋgər] 'slant'

(369) /gorgəl/ ——> [górgəl] 'throat'

(370) /kəndiy+ə/ ——-> [kə́ndiyə] 'oil lamp'

(371) /jinjiwir+i/ ——> [ǰíñǰiviri] 'ginger'

ii Words with underlying long vowels in their initial syllables: the initial syllable vowel is stressed and long. This class consists of the majority of monosyllables and vowel final disyllabic nouns, adjectives

and adverbs.

(372) /o:y/ ——> [ó:y] 'eye' N

(373) /pa:m/ ——> [pá:m] 'bread'

(374) /pa:y/ ——> [pá:y] 'father'

(375) /sa:l/ ——> [sá:l̥] 'salt'

(376) /ka:ws/ ——> [ká:ws] 'matter' N

(377) /le:/ ——> [lé:] 'religion'

(378) /wi:/ ——> [ví:] 'come'

(379) /fi:w/ ——> [fí:w] 'banana, plantain'

(380) /cu:w+a/ ——> [čú:wə] 'rain' N

(381) /ba:k+ə/ ——> [bá:kə] 'cow, bull'

(382) /be:s+u/ ——> [bé:su] 'lip'

(383) /se:w+i/ ——> [sé:vi] 'fence'

(384) /li:m+ə/ ——> [lí:mə] 'file' N

(385) /ma:c+i/ ——> [má:či] 'male'

(386) /si:nz+ə/ ——> [sí:nzə] 'ash'

(387) /tri:y+ə/ ——> [trí:yə] 'bring'

(388) /fru:$^{w}/_{g}$+ə/ ——> [frú:wə ~ frú:gə] 'boil' V_{int}

(389) /pi:ris/ ——> [pí:ris] 'saucer'

(390) /wa:rzim/ ——> [vá:rzim] 'paddy field, cultivation'

(391) /a:nim+ə/ ——> [á:n̥imə] 'courage'

(392) /glɔ:riy+ə/ ——> [gló:riyə] 'heaven; glorious'

(393) /æ:tik+ə/ ——> [ǽ:tikə] 'tuberculosis'

(394) /gɔ:mitu/ ——> [gó:mitu] 'vomit'

(395) /lu:wriy+ə/ ——> [lú:wriyə] 'rent' N

iii Words with underlying long vowels in non-initial syllables: the long vowel is stressed and potentially half-long. This class can be further divided into two subgroups depending on whether the long vowel is in the final or penultimate syllable (only these two types occur). The first subgroup includes most consonant final polysyllables, polysyllabic verbs in -i: and -a: (historically from SP verbs in -ir and -ar), and a few others. The second subclass contains most polysyllabic words not in the first subclass: i.e. vowel final non-verbs and verbs in -ə.

(396) /ənima:l/ ——> [ənimá˙l ~ ənimál] 'animal'

(397) /nəri:s/ ——> [nərí˙s ~ nərís] 'nose'

(398) /kəmi:m/ ——> [kəmí˙m ~ kəmím] 'way, path'

(399) /kəsti:w/ ——> [kəstí˙w ~ kəstíw] 'difficulty'

(400) /əruməsa:m/ ——> [əruməsá˙m ~ əruməsám] 'roof'

(401) /kəkərla:t/ ——> [kəkərlá˙t ~ kəkərlát] 'cockroach'

(402) /əwr+i:/ ——> [əwrí˙ ~ əwrí] 'open' V

(403) /kim+a:/ ——> [kimá˙ ~ kimá] 'burn'

(404) /kuns+a:/ ——> [kunsá˙ ~ kunsá] 'begin'

(405) /enkontr+a:/ ——> [eŋkontrá˙ ~ eŋkontrá] 'meet'

(406) /pəpiy+a:/ ——> [pəpiyá˙ ~ pəpiyá] 'speak, talk'

(407) /mers+e:/ ——> [mɜrsé˙ ~ mɜrsé] 'thanks' N

(408) /kədi:y+ə/ ——> [kədí˙yə ~ kədíyə] 'chain'

(409) /konte:nt+i/ ——> [konté˙nti ~ konténti] 'happy'

(410) /iski:s+ə/ ——> [iskí˙sə ~ iskísə] 'forget'

(411) /usku:nd+ə/ ——> [uskú˙ndə ~ uskúndə] 'hide'

(412) /risi:b+ə/ ——> [risí˙bə ~ risíbə] 'receive'

Many alternations between these patterns can be observed. In (413) - (419) alternations between type i and type iii are produced by the addition of a long vowel suffix to a stem containing no long vowels.

(413) a) /tis+ə/ ——> [tísə] 'rattan' V

b) /tis+ə+sa:m/ ——> [tisəsá˙m ~ tisəsám]
'rattan work'

(414) a) /inc+ə/ ——> [íñčə] 'fill'

b) /inc+i:+d+u/ ——> [iñčí˙du ~ iñčíḍu]
'filled'

(415) a) /prend+ə/ ——> [préndə] 'learn'

b) /prend+do:r/ ——> [prendudó˙r ~ prendudór][19]
'learned one'

(416) a) /ker+ə/ ——> [kérə] 'want'

b) /ker+i:y+ə/ ——> [kerí˙yə ~ keríyə]
'desire, purpose'

(417) a) /perd+ə/ ——> [pɜ́rdə] 'lose'

b) /perd+i+sa:m/ ——> [pɜrdisá˙m ~ pɜrdisám]
'loss'

(418) a) /səw+ə/ ——> [sə́və] 'know'

b) /səw+i:+nd+u/ ——> [səví˙ndu ~ səvíndu]
'knowing'

---

[19]For epenthetic [u], see rule (510) TBEp below.

(419) a) /kum+ə/ ——> [kúmə] 'eat'

b) /kum+e:r+ə/ ——> [kumé˙rə ~ kuméra] 'food'

c) /kum+i:+d+u/ ——> [kumí˙du ~ kumídu] 'eaten'

In these examples the stems contain only short vowels. The same stress alternations can be produced by the addition of a long vowel suffix to a stem containing a long vowel, since when a word has two or more underlying long vowels, it is the last of these which is stressed. But here length alternations in the stem vowel are involved as well, because, as mentioned above, unstressed vowels must be short. Thus the alternation is between types ii and iii. Examples (420) - (430) illustrate the situation.

(420) a) /mi:d+ə/ ——> [mí:də] 'measure'[20]

b) /mi:d+i:/ ——> [midí˙ ~ midí] 'measure' V

c) /mi:d+i:+do:r/ ——> [mididó˙r ~ mididór] 'surveyor'

(421) a) /owdi:s+ə/ ——> [owdí˙sə ~ owdísə] 'obey'

b) /owdi:s+i:+d+u/ ——> [owdisí˙du ~ owdisídu] 'obedient'

---

[20] A standard unit of capacity equivalent to about forty ounces.

(422) a) /wɜ:rd+i/ ——> [vɜ́:rdi] 'green'

b) /wɜ:rd+u:r+ə/ ——> [vɜrdú˙rə ~ vɜrdúrə]
'vegetable'

(423) a) /pene:r+ə/ ——> [pené˙rə ~ penérə]
'sieve'

b) /pene:r+a:/ ——> [penerá˙ ~ penerá]
'sift'

(424) a) /sæ:t+i/ ——> [sǽ:ti] 'seven'

b) /sæ:t+æ:nt+ə/ ——> [sætǽ˙ntə ~ sætǽntə]
'seventy'

(425) a) /fu:l+ə/ ——> [fú:lə] 'flower'

b) /fu:l+e:r+ə/ ——> [fulé˙rə ~ fulérə]
'flowering tree'

(426) a) /usku:nd+ə/ ——> [uskú˙ndə ~ uskúndə]
'hide'

b) /usku:nd+i:+d+u/ ——> [uskundí˙du ~ uskundídu]
'hidden'

(427) a) /do:d+u/ ——> [dó:du] 'stupid'

b) /do:d+i:s+ə/ ——> [dodí˙sə ~ dodísə]
'stupidity'

(428) a) /əmo:r/ ——> [əmó˙r ~ əmór] 'love' N

b) /əmo:r+o:z+u/ ——> [əmoró˙zu ~ əmorózu]
'loving'

(429) a) /nɔ:w+i/ ——> [nɔ́:vi] 'nine'

b) /nɔ:w+æ:nt+ə/ ——> [nɔvǽ˙ntə ~ nɔvǽntə]
'ninety'

(430) a) /ba:rw+ə/ ——> [bá:rvə] 'beard'

b) /ba:rw+iy+a:/ ——> [bərviyá· ~ bərviyá] 'shave'

c) /ba:rw+e:r+u/ ——> [bərvé·ru ~ bərvéru] 'barber'

To account for these length alternations[21] we need the following rule, which will apply after the Main Stress Rule to shorten all unstressed long vowels of a word:

(431) Length Loss (LL):

$$\left[\begin{matrix} V \\ -\text{stress} \end{matrix}\right] \longrightarrow [-\text{long}]$$

This completes the broad outlines of stress and vowel length at the word level in BP. Before going on to consider additional data, it is a good idea to pause and briefly recapitulate. It has been shown that vowel length must be part of the phonological representation of a word and that stress is predictable. We have formulated three rules: a Main Stress Rule, which places stress on the last long vowel of a word, a rule of Length Loss, whereby all unstressed underlying long vowels become surface short vowels, and a scalar rule of Length Adjustment, whereby stressed long vowels in non-initial syllables are reduced in length. The operation of these rules is illustrated in

---

[21] The quality alternation between [a:] and [ə] in (430) can be accounted for by a later reduction rule; see (603) Raising.

the derivations of (420) a), b),and c):[22]

(432)

/mi:d+ə/	/mi:d+i:/	/mi:d+i:+do:r/	
í:	í:	ó:	MSR (rule (343))
	i	i i	LL (rule (431))
_____	_____i(˙)_	_____o(˙)_	LA (rule (342))
[mí:də]	[midí(˙)]	[mididó(˙)r]	

In order to simplify future phonetic transcriptions, stressed vowels in non-initial syllables will be indicated as half-long only.

In the above we have been talking only of primary stress. Any word may have an optional tertiary stress preceding its main stress (unless this falls on the first syllable). The position of the optional [3stress] depends on the syllable structure of the part of the word preceding the main stress. Tertiary stress may fall on the syllable immediately preceding main stress only when this falls on the second syllable. More usually in this case no [3stress] occurs: the language appears to be antipathetic towards two succeeding stressed syllables.

(433) /sənt+a:/ ——> [sənt$\overset{1}{\grave{a}}$˙ ~ s$\overset{3}{\grave{ə}}$nt$\overset{1}{\grave{a}}$˙] 'sit'

(434) /nəta:l/ ——> [nət$\overset{1}{\grave{a}}$˙l ~ n$\overset{3}{\grave{ə}}$tt$\overset{1}{\grave{a}}$˙l] 'Christmas'

(435) /iñæ:m+ə/ ——> [iñ$\overset{1}{\grave{æ}}$˙mə ~ iñ$\overset{3}{\grave{æ}}$˙mə] 'manioc'

When two syllables precede the main stress, the initial

[22] As mentioned earlier, for BP, it is not necessary to assume rule cycles which apply to stretches below the level of the word. Thus each of the three derivations in (432) contains only one rule cycle.

syllable may receive [3stress].

(436) /ka:r+əstrí:y+ə/ ——> [kərəstrí^{1}·yə ~ kə́rəstrí^{1}·yə]
'expense' (?)

(437) /kəkərla:t/ ——> [kəkərlá^{1}·t ~ kə́kərlá^{1}·t]
'cockroach'

(438) /enkontr+a:/ ——> [eŋkontrá^{1}· ~ éŋkontrá^{1}·] 'meet'

(439) /kərdəmu:ng+u/ ——> [kərdəmú^{1}·ŋgu ~ kə́rdəmú^{1}·ŋgu]
'cardamom'

(440) /mi:d+i:+sa:m/ ——> [midisá^{1}·m ~ mídisá^{1}·m]
'measurement'

When three or more syllables precede the main stress, [3stress] may fall on the second syllable if it is closed, otherwise on the initial syllable.

(441) /əwiyəmæ:nt+u/ ——> [əviyəmǽ^{1}·ntu ~ ə́viyəmǽ^{1}·ntu]
'tool'

(442) /inti:nd+mæ:nt+u/ ——>
[intindumǽ^{1}·ntu ~ intíndumǽ^{1}·ntu][23]
'meaning'

(443) /trəwiya:+do:r/ ——> [trəviyədó^{1}·r ~ trə́viyədó^{1}·r]
'worker'

(444) /əruməsa:m/ ——> [əruməsá^{1}·m ~ ə́ruməsá^{1}·m] 'roof'

We may restate the distribution as follows: tertiary stress may fall on the second syllable, provided it is closed and that at least one syllable intervenes before the

[23]For epenthetic [u], see rule (510) TBEp below.

primary stress; otherwise [3stress] falls on the first syllable (provided this is not the position of primary stress). This is stated formally in (445).

(445) Tertiary Stress Assignment (TSA): optional

$$\left[{}^{V}_{-\text{stress}}\right] \longrightarrow [3\text{stress}]$$

$$/\left\langle \left[_{L}\ C_{o}\left[{}^{V}_{-\text{stress}}\right]\right\rangle C_{o}-\left\langle \left[{}^{C}_{\alpha\text{son}}\right]\left[{}^{C}_{\beta\text{nas}\ \gamma\text{son}}\right]C_{o}\left[{}^{V}_{-\text{stress}}\right]\right\rangle$$

if $\alpha = -$, then $\beta = \gamma$

L is a variable ranging over all the lexical categories.

The condition on α, β, and γ captures the fact that an intervocalic cluster of obstruent plus liquid is assigned to the syllable of the following vowel: see Section 2 on syllable structure and cf. (447) below. The following stress patterns result from the operation of (343) and (445) on related forms.

(446) a) /owdi:s+ə/ ——> [$\overset{(3)}{o}$wdi$\overset{1}{\cdot}$ssə] 'obey'

b) /diz+owdi:s+ə/ ——> [d$\overset{(3)}{i}$zowdi$\overset{1}{\cdot}$ssə] 'disobey'

c) /owdi:s+i:+d+u/ ——> [$\overset{(3)}{o}$wdisi$\overset{1}{\cdot}$du] 'obedient'

d) /diz+owdi:s+i:+d+u/ ——> [diz$\overset{(3)}{o}$wdisi$\overset{1}{\cdot}$du] 'disobedient'

(447) a) /ələgr+i:y+ə/ ——> [$\overset{(3)}{ə}$ləgri$\overset{1}{\cdot}$yə] 'happiness'

b) /ələgr+i:y+ə+me:nt+i/ ——> [$\overset{(3)}{ə}$ləgriyəme$\overset{1}{\cdot}$nti] 'happily'

(448) a) /inti:nd+ə/ ——> [$^{(3)}$ɪ^{1}ntɪ^{1}·ndə] 'understand'

b) /inti:nd+i:+d+u/ ——> [$^{(3)}$ɪ^{1}ntind$^{(1)}$·du]

'understood'

c) /inti:nd+mæ:nt+u/ ——> [int$^{(3)}$ɪ^{1}ndumæ^{1}·ntu]

'meaning'

Note that the vowel in the second syllable of (448) c) is shortened from /i:/ to [i]. In order for this process not to be blocked, LL, which applies only to [-stress] vowels, must be ordered before TSA.

There appear to be at most four stress levels in BP; i.e. [4stress] is equivalent to [-stress]. In syntactic concatenations, the [3stress] assigned by TSA in automatically reduced by the Nuclear Stress Rule (460) below to [4stress]. Thus TSA is a phenomenon of the citation form. It will be largely ignored in future transcriptions.

There is a group of verbal prefixes denoting tense/aspect which pose a problem for our analysis of stress and vowel length. These are illustrated in (449) - (451).

(449) a) [té:m] 'is'

b) [nunté·m] 'is not'

c) [loté·m] 'will be'

(450) a) [dá:] 'give'

b) [tədá·] 'is giving'

c) [nukudá·] 'is not giving'

(451) a) [bʒvə] 'drink'

b) [təbʒvə] 'is drinking'

c) [ǰabʒvə ~ ǰəbʒvə] 'drank'

Note first of all that LA has applied to the stressed vowels in (449) b), (449) c), (450) b), etc. leaving tham half-long. It can be demonstrated that LA is a word level rule: it does not, for example, apply to /ɔ:r+ə/ in (452)

(452) [dɒ̄^{2}:z ɔ̄^{1}:rə] 'two o'clock'

Consequently the forms in b) and c) of (449) - (451) must be treated as inflected verbs; i.e. as single words rather than phrases. This analysis is consistent with the distributional evidence: the tense/aspect markers exhibited above occur only preceding a verb stem, from which they may be separated by at most one other such marker. In this case, however, MSR as formulated in (343) should assign stress to the prefix in (451) b) and c). Furthermore, the alternation between [ə] and [a] in (451) c) seems to indicate that LL has applied to underlying /ja:-/ with a later optional reduction rule producing [ə]. This behavior is different from that already observed in forms such as (443), where reduction is obligatory. It is generally recognized that non-productive derivational affixes (which are part of the lexical representation of the word) and productive affixes (which are transformationally introduced) may exhibit different phonological behavior. Chomsky and

Halle[24] have pointed out for example that such affixes as -ing in English are stress neutral and that phonological processes such as /g/ ——> Ø / [+nas]_#, and [+son] ——> [+syll] / C_#, which normally take place before a word boundary (but not before the morpheme boundary +) also take place before these affixes. Their solution, which we shall adopt here, is to posit a word boundary between the stem and the transformationally introduced affix. To be consistent, as well as having such underlying representations as /nun#te:m/ (NEG-is) and /ja:#bew+ə/ (PAS-drink) for (449) b) and (451) c), with a word boundary between stem and inflectional prefix, we must also have /di:y+ə#ntu/ (day-LOC) and /kɔ:nt+ə#pə/ (amount-DAT) for (453) b) and (454) b) with a word boundary between stem and inflectional suffix.

(453) a) [dí:yə] 'day'

b) [dí:yəntu] 'in the day'

(454) a) [kɔ́:ntə] 'amount'

b) [kɔ́:ntəpə] 'to an amount'

There is no way to reformulate MSR to apply to such strings and produce [jabʒ́və] 'drank', but [dí:yəntu] 'on the day', without recourse to a cyclic analysis. Words containing transformationally introduced affixes will, therefore, have underlying representations of the following:

---

[24]SPE pp. 84 - 86.

type: /[nun[te:m]$_V$]$_V$/, /[ja:[bew+ə]$_V$]$_V$/, /[[di:y+ə]$_N$ntu]$_N$/, /[[kɔ:nt+ə]$_N$pə]$_N$/. The rules for stress and vowel length will then apply first to the strings within the internal brackets. These brackets are then erased (leaving a word boundary # between the stem and the affix, since the stem is dominated by a syntactic category node),[25] and the cycle of rules reapplies to the whole string. Since phonological rules cannot ignore word boundaries,[26] MSR as formulated in (343) will not be able to apply on the second cycle to a prefix separated from the stem by a # boundary. It would, however, stress a suffix containing a long vowel, as X in (343) could contain #. But as all unstressed suffixes encountered so far contain only underlying short vowels,[27] no modification to (343) is necessary at this point.

Length Adjustment will have to be modified to allow for a word boundary between the vowel to be shortened and the previous syllabic element in the environment. It will not, however, apply in syntactic strings such as (452), which contain a double word boundary.

(455) Length Adjustment (LA): final version

$$V \longrightarrow [nlong] \;/\; V\ C_o(\#)C_o\left[\underline{\quad}_{1stress}\right]$$

$$n = 0 \text{ or } 1$$

---

[25] All boundaries are erased as the final step in any derivation.

[26] Formative (+) boundaries may, of course, be ignored.

[27] See (489) - (495) for a class of exceptions.

Length Loss will apply in its present form, as will TSA. Some derivations are given below.[28]

(456) /[ja:[bew+ə]$_V$]$_V$/ (PAS-drink)

1	MSR (rule (343))
	LL, TSA
———	
	MSR
a	LL (rule (431)
a/ə ɜv	Other rules
(3)	TSA
———	
[ǰ$^{(3)}$ab^{1}ɜvə ~	'drank'
ǰ$^{(3)}$əb^{1}ɜvə]	

(457) /[nun[te:m]$_V$]$_V$/ (NEG-be)

1	MSR
	LL, TSA
———	
	MSR
	LL
(3)	TSA
e·	LA
———	
[n$^{(3)}$unt^{1}e·m]	'isn't'

---

[28] In accordance with usual practice, the replacing of brackets by word boundaries, and the final erasure of the latter are not explicitly shown. Derivational cycles are separated by horizontal lines. Low level rules, such as LA, are assumed to be post-cyclic.

(458) /[[di:y+ə]$_N$ntu]$_N$/ (day-DAT)

1	MSR
	LL, TSA,
	MSR, LL, TSA, LA
[dī^1:yəntu]	'on the day'

(459) /[[me+di:y+ə]$_N$pə]$_N$/ (midday-DAT)

1	MSR
	LL
(3)	TSA
	MSR, LL, TSA
i·	LA
[m$^{(3)}$ē^{1}dī·yəpə]	'at midday'

In the above discussion I alluded to the fact that the phonological processes applicable at the word level differ from those applicable to larger constructions. By way of illustration, I shall briefly describe stress and vowel length in simple syntactic constructions, before returning to a treatment of compounds, which seem to form an intermediate type.

Rule (460), an analog to the Nuclear Stress Rule proposed by Chomsky and Halle for English, cyclically assigns stress to the last primary stressed vowel in a syntactic construction.[29] As is the case in English, each assignment

[29] Cf. SPE p. 91. It is not immediately clear which nodes are cyclic in BP. We shall assume that NSR applies at least to NP, VP, and S.

of primary stress to a string decreases the stress of all the other vowels in the string by one degree. Thus we have the stress patterns of (452) above and (461) - (462) below.

(460) Nuclear Stress Rule (NSR):

$$V \longrightarrow [1stress] \; / \; \left[\frac{\quad}{1stress}\right] X]_{NP\ VP\ S}$$

where X contains no instance of [1stress]

(461) /[[[mæ:z+ə]$_N$su]$_N$[pa:n+u]$_N$]$_{NP}$/ ——> [m$\overset{2}{æ}$:zəsu p$\overset{1}{a}$:ṇu]

(table GEN cloth) 'tablecloth'

(462) /[[pa:n+u]$_N$[bist+i:]$_V$]$_S$/ ——> [p$\overset{2}{a}$:ṇu bist$\overset{1}{i}$·]

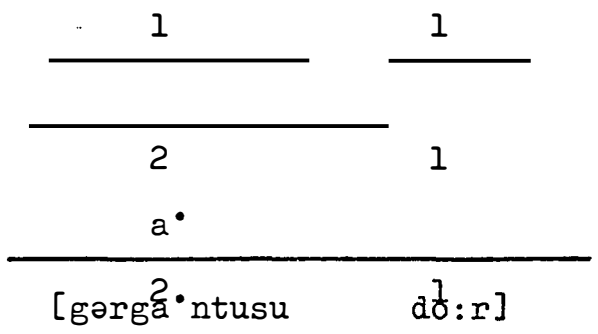
(cloth wear) 'wear a cloth (i.e. sari or dhoti)'

As LA does not apply across a double word boundary the long vowel in /pa:n+u/ does not get shortened in (461). Furthermore, LL shortens only unstressed vowels and hence will not apply to the [2stress] vowels in (461) and (462). Some full derivations are illustrated below.

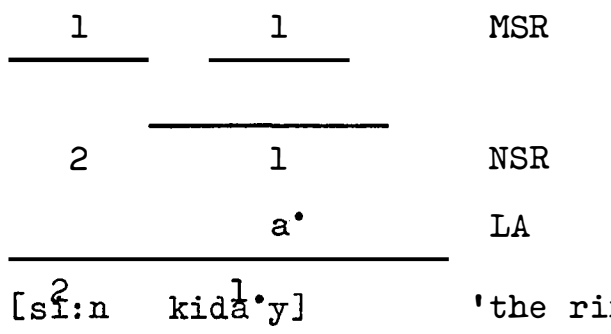
(463) /[[ma:l]$_A$[tæ:mp+u]$_N$]$_{NP}$/ (bad time[s])

1	1	MSR (rule (343))
ḷ		Other rules [30]
2	1	NSR
[m$\overset{2}{a}$:ḷ	t$\overset{1}{æ}$:mpu]	'bad times'

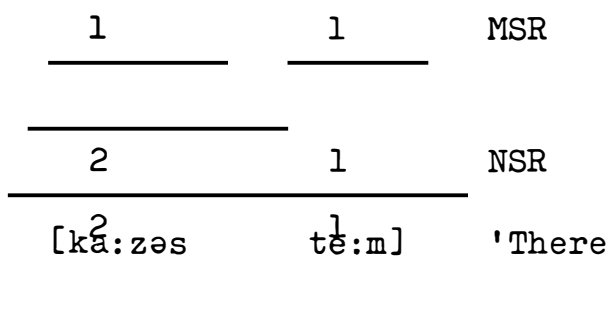
(464) /[[[gərga:nt+u]$_N$su]$_N$[do:r]$_N$]$_{NP}$/ (throat-GEN pain)



(465) /[[si:n]$_N$[ki[da:y]$_V$]$_V$]$_{VP}$/ (bell[s] DESC-ring)



(466) /[[[ka:z+ə]$_N$s]$_N$[te:m]$_V$]$_S$/ (house-PL are)



(467) /[[tə[tir+a:]$_V$]$_V$[[ɔ:r+ə]$_N$s]$_N$]$_{NP}$/ (PRES-take hour-PL)

1	1	MSR
2	1	NSR
a·		LA
[tətirā· ɔ:rəs]		'when taking'

A sequence of two stressed syllables such as that in (467) would be obtained only in very careful speech. In more rapid, casual style, (468) or (469) would be more likely.

(468) [tətí^{2}ra ɔ:rəs]

(469) [tətí^{2}rə ɔ:rəs]

Note that the destressed vowel is always short. The optional reduction of underlying /a:/, which during the derivation was stressed by MSR at the word level (and hence not subject to automatic raising to [ə] by (603)) but later destressed by phrase-level rules, is undoubtedly the same phenomenon as that found above with respect to tense/aspect affixes: cf. (451) c). We are unable to give a full description of the stress shifting phenomenon exhibited here. It is not clear for example how close the two stressed syllables must be before shifting becomes possible: (470) - (471) show that they need not be adjacent.

(470) /[[ja:[tom+a:]$_V$]$_V$[pesa:m]$_N$]$_{NP}$/ ——>

(PÁS buy person)

[jət$\overset{2}{o}$ma pes$\overset{1}{a}$·m] 'the person who bought [it]'

(471) /[[[iruma:m]$_N$su]$_N$[fi:y+u]$_N$]$_{NP}$/ ——>

(sibling GEN son)

[$\overset{2}{i}$rumansu f$\overset{1}{i}$:yu][31] 'nephew'

When the word with the shifting stress has at least two pretonic syllables, stress shifts to the second syllable if it is closed, otherwise stress shifts to the first syllable.

(472) a) /əsərt+a:/ ——> [$\overset{(3)}{ə}$sərt$\overset{1}{a}$·] 'meet'

b) /[[tə [əsərt+a:]$_V$]$_V$[[ɔ:r+ə]$_N$s]$_N$]$_{NP}$/ ——>

(PRES meet hour PL

[təəs$\overset{2}{ə}$rtə $\overset{1}{ɔ}$:rəs] 'when meeting'

The following rule is tentatively proposed.[32]

---

[31] In this example unstressed /a:/ has not reduced to [ə], though presumably reduction is possible here too. It seems that the phrase-level rule which reduces destressed a to [ə] (see (474)) operates more frequently before a word boundary.

[32] The rule as written does not account for the stress shifting which may occur after NSR has applied on the next higher cycle: instead of the expected stress pattern [...3...2...1...], we often find [...2...(3) ..1...] in all but very careful speech.

E.g. /[[[do:z]$_A$[[ɔ:r+ə]$_N$pə]$_N$]$_{NP}$[ka:rt+u]$_N$]$_{NP}$/

(two hour DAT quarter)

[d$\overset{3}{o}$:z $\overset{2}{ɔ}$:rəpə k$\overset{1}{a}$:rtu] ~ [d$\overset{2}{o}$:z $\overset{3}{ɔ}$:rəpə k$\overset{1}{a}$:rtu] ~

[d$\overset{2}{o}$:z ɔrəpə k$\overset{1}{a}$:rtu] 'quarter to two'

We assume that this and SS are closely related processes,

(473) Stress Shift (SS): tentative formulation

allegro speech

SD: $\#\langle CV\rangle C_o V \left\langle \begin{bmatrix} C \\ \alpha son \end{bmatrix} \begin{bmatrix} C \\ \beta nas \\ \gamma son \end{bmatrix} \right\rangle X \begin{bmatrix} V \\ mstress \end{bmatrix} C_o \# Y \begin{bmatrix} V \\ +stress \end{bmatrix}$

1 2 3 4 5 6 7 8 9 10 11

Conditions: a) if α = - then $\beta = \gamma$

b) unknown conditions on X and Y

SC: 4 ——> [mstress]

7 ——> [-stress]

After (473), LL followed by (474) will apply. Some derivations follow.

(474) A Reduction (ARed): optional

$$\begin{bmatrix} V \\ +back \\ -round \\ -stress \end{bmatrix} \longrightarrow [-low]$$

Note: especially common / _ #

(475) /[[enkulk+a:]$_V$[oy+a:]$_V$]$_{VP}$/ (enquire see)

	1	1	MSR (rule (343))
ŋ			Other rules
	2	1	NSR (rule (460))
2	4		SS
	a		LL (rule (431))
	(ə)		ARed (optional)
		ə·	LA (rule (455))
[eŋkulk^{2}a/ə		oyə^{1}·]	'find out'

though whether both should be accounted for by the same rule will be left an open question.

(476) /[[pikini:m]$_A$[ka:z+ə]$_N$]$_{NP}$/ (small house)

	1	1	MSR
	2	1	NSR
2	4		SS
	i		LL
	ŋ		Other rules
[^{2}pikiniŋ		^{1}kà:zə]	'small house'

(477) /[[tə[fik+a:]$_V$]$_V$[o:r+u]$_N$]$_{NP}$/ (PRES-become gold)

	1	1	MSR
	2	1	NSR
2	4		SS
	a		LL
	(ə)		ARed
[tə^{2}fik^{a}⁄ə		^{1}ò:ru]	'becomes gold'

We shall now return to a consideration of compounds, which, as can be seen from (478) - (482), are similar to syntactic strings with respect to primary stress placement, but like simple words with respect to length phenomena.

(478) a) /sa:l/ ——> [sá:l] 'salt'

b) /a:w/ ——> [á:w] 'water'

c) [$^{(3)}$sa^{1}là·w ~ sə$^{(3)}$^{1}là·w] 'saltwater'

(479) a) /di:y+ə/ ——> [dí:yə] 'day'

b) /gra:nd+i/ ——> [grá:ndi] 'big'

c) [dìyə(g)rá·ndi] 'the Christmas to New Year's holiday'

(480) a) /bɔ:wr+ə/ ——> [bɔ́:vrə] 'pumpkin'

b) /də/ ——> [də] 'of'

c) /a:w/ ——> [á:w] 'water'

d) [bɔvərdá·w] 'a kind of watermelon' = BT /na:Tanka:y/

(481) a) /sæ:t+æ:nt+ə/ ——> [sætǽ·ntə] 'seventy'

b) /si:nk+u/ ——> [sí:ŋku] 'five'

c) [sætæ̀ntəsí·ŋku] 'seventy-five'

(482) a) /wi:nt+i/ ——> [ví:nti] 'twenty'

b) /[[ung]$_A$ə]$_N$/ ——> [úŋŋə] 'one'[33]

c) [vìntəúŋŋə] 'twenty-one'

It is clear that the last element of a compound receives primary stress, even if, as in (482) c), it contains no long vowel. However, the non-primary stressed vowels are short, with at most tertiary stress, and the primary stressed vowel undergoes Length Adjustment. There are several possible ways to account for these facts. The two major alternatives are:

1. Assume an underlying representation such as /[sa:l#a:w]$_N$/ and modify MSR so as to assign stress to the last member of compounds. All other rules will apply

[33]This derivation will be discussed in detail on pp. 136 ff.

as formulated.

2. Assume underlying /[[sa:l]$_N$[a:w]$_N$]$_N$/. On the first cycle MSR will assign [1stress] to each of the two constituents. NSR could then easily be modified to apply on the word level to produce a [2...1] stress pattern on the next cycle. The [2stress] could be erased by an ad hoc rule and other rules would apply as above.

Neither of these solutions is appealing. We shall choose the former because the representation seems to best reflect the idiosyncratic nature of compounds.[34] The new formulation of MSR appears in (483); it is designed to assign [1stress] to the last long vowel (or the first short vowel if no long vowel is present) in that part of the lexical item which follows the last internal word boundary, provided stress has not already been assigned to the part of the word preceding the internal word boundary on an earlier cycle. This last proviso is necessary to prevent the rule from acting on the stress neutral suffixes discussed earlier.

(483) Main Stress Rule (MSR): final version

$$V \longrightarrow [1\text{stress}]$$

$$/\ [\ (Y\#)\ \langle X\rangle\ C_o \left[\overline{\langle +1\text{ong}\rangle}\right]\ (C_o\left[\overset{V}{-1\text{ong}}\right]C_o)_o]_L$$

where L ranges over the lexical categories and

Y contains no [1stress]

There will be some asymmetry in the operation of (483): it

[34] Note, for example, the unpredictable elipsis of /g/ in (479) c) and the epenthesis of /ə/ in (480) d).

will vacuously re-apply on the second cycle of /[ja:[bew+ə]$_V$]$_V$/ but will be blocked from re-applying on the second cycle of /[[di:y+ə]$_N$pə]$_N$/. Some derivations follow.

(484) /wi:nt+ə#do:z/ (twenty-two)

	l	MSR
i		LL (rule(431))
v	s	Other rules
(3)		TSA
	o•	LA (rule (455))

[v$^{(3)}$ɪ$_{1}$ntədo•s] 'twenty-two'

(485) /te:#a:w/ (tea water)

	l	MSR
e		LL
y		Other rules
(3)		TSA
	a•	LA

[t$^{(3)}$e$_{1}$ya•w] 'tea (the liquid)' = BT /te:+taNNi/

(486) /pa:n+u#wæ:y/ (cloth-old)

	l	MSR
a		LL
n•	v	Other rules
ə		ARed
(3)		TSA
	æ•	LA

[p$^{(3)}$ə$_{1}$nuvæ•y] 'rag'

(487) /po:k#ida:y/ (little-age)

1	MSR
o	LL
(3)	TSA
a•	LA

[p(3)o kidă•y] 'young'

(488) /bɔ:wər#d#a:w/ (pumpkin-of-water)

1	MSR
ɔ	LL
v	Other rules
(3)	TSA
a•	LA

[b(3)ɔ vərdă•w] 'watermelon'

This basically completes our analysis of stress and length in BP. There is, however, one small group of items which have not been accounted for, and which will require exceptional treatment. The items, given in (489) - (495) are all [+human] nouns[35] which end in unstressed unreduced [a].

(489) [bú:wa ~ bú:] 'elder brother' = BT /aNNan/; cf. BP [čú:wə] 'rain'

(490) [číča ~ číči] 'elder sister' = BT /akka:/; cf. BP [či:čə] 'gekko'

[35]None of these words is derived from SP.

(491) [pó̤diyas ~ pó̤di] 'child' cf. S /poDi/ 'small'
BT /poTiyan/ ([pó̤diyən]) 'small boy'

(492) [grǽnpəpa ~ grǽnpa] 'grandfather' < E?

(493) [grǽnməma ~ grǽnma] 'grandmother' <E?

(494) [pó̤pa] 'father' < E?

(495) [mǽma] 'mother' < E?

It seems clear that the unreduced [a]'s will have to come from underlying long /a:/'s. Two problems then arise: if the underlying representations are to be /bu:wa:/ etc., MSR as formulated will stress the long /a:/, and LL and LA will apply yielding *[buwá• ~ buwá]; moreover, even if stress can be correctly placed on /u:/, the final unstressed /a:/ will have to be prevented from undergoing optional A Reduction (rule (474)).

Stress will be correctly assigned to these words if we assume underlying /[[bu:]$_N$a:]$_N$/,[36] /[[cic]$_N$a:]$_N$/, etc., and the final vowels will not undergo Reduction if we simply mark these items as exceptions to (474). This analysis is admittedly ad hoc: the word boundary, for example, has no independent justification, (though the variants in (489) - (491)[37] could be used to support at least a + boundary). I feel, however, that as the number of items in this group is so small, their exceptional nature should as much as possible be accounted for by their lexical representations,

[36] [w] will be inserted by Glide Epenthesis (rule (535)).

[37] There appears to be no semantic or distributional difference between any of these variants.

rather than by a major overhaul of existing rules.

4 Vowels

Vowel quantity has already been discussed at length in the previous section. Here we shall be concerned with other vowel related phenomena: epenthesis, elision, and quality variations.

4.1 Epenthesis

An optional epenthetic vowel occurs between a stop or spirant and a following /r/ when these are part of an intervocalic cluster preceded by a stressed vowel. The epenthetic vowel has the quality of the following unstressed vowel (only /i u a/ occur).

(496) Epenthesis in Clusters (ClEp):

optional

SD: $\begin{bmatrix} V \\ +stress \end{bmatrix}_1 \begin{bmatrix} C \\ -son \end{bmatrix}_2 \begin{bmatrix} C \\ +son \\ -nas \\ -lat \end{bmatrix}_3 V_4$

SC: Copy 4 between 2 and 3.

(497) a) /ka:mbr+ə/ ——> [ká:mbərə ~ ká:mbrə] 'room'

b) /ka:mbr+a:d+u/ ——> [kəmbrá'du] 'friend'

(498) a) /so:gr+u/ ——> [só:guru ~ só:gru] 'father-in-law'

b) /sɔ:gr+ə/ ——> [só:gərə ~ só:grə] 'mother-in-law'

(499) a) /bu:fr+ə/ ——> [bú:fərə ~ bú:frə] '(water) buffalo'

(500) a) /ka:tr+u/ ——> [ká:turu ~ ká:tru] 'four'

b) /ka:tr+e:r/ ——> [kətré·ᵊr] 'fourth'

(501) a) /pa:dr+i/ ——> [pá:diri ~ pá:dri] 'priest'

There is no contrast between sequences of the types in (502) when 3 is [-son].

(502) a) $\underset{1}{[V_{\alpha stress}]}(C)\underset{2}{C}\underset{3}{r}\underset{5}{[V_{-\alpha stress}]}$

b) $\underset{1}{[V_{\alpha stress}]}(C)\underset{2}{C}\underset{3}{V}\underset{4}{r}\underset{5}{[V_{-\alpha stress}]}$

In the cluster and epenthesis analysis proposed here, type a) is taken to be basic and b) derived. Alternatively, type b) could be viewed as underlying in the above examples; a rule would be necessary to delete the second vowel of such a sequence obligatorily when the third vowel is stressed, and optionally otherwise. The fact that in all such sequences the second and third vowels must be identical for the second to appear on the surface would not be explained by such an analysis, and underlying C<u>r</u> clusters would still be needed in initial position and in non-alternating forms such as (503).

(503) /ələ<u>gr</u>+i:y+ə/ ——> [ələ<u>gr</u>í·yə] 'happiness'

Historically, the only natural way to account for epenthetic /b/ in (497), is to posit a cluster at some stage, i.e. to assume the development was something like

SP câmara /kámara/ > /kamra/ > /kambra/ > /ka:mbra/.[38] Cf. the similar development of Latin humerus to SP hombro /ombro/. The /t/ in (504) can be explained in the same way.

(504) /pa:stru/ ——> [pá:sturu ~ pá:stru] 'bird'
Cf. SP pássaro /pásaru/.

ClEp does not apply when 3 in (502) is sonorant.

(505) /roza:yr+u/ ——> [rozá˙yru] 'rosary'
(506) /a:lwr+i/ ——> [á:lvri] 'tree' Cf. SP árvore.
(507) /jæ:nr+u/ ——> [ǰǽ:NDru] 'son-in-law' Cf. SP genro.

The symbols [N D] in the last example represent consonants with retracted (but not retroflex) apical articulation. The epenthetic stop [D] is added by a later rule which also applies across word boundaries (see rule (742) DEp): doubtless a similar process to that which gave rise to epenthetic /b/ and /t/ discussed above. In this case, however, the rule is still synchronic. In (506) [v] is produced by the optional Glide Constriction rule (705), which must be ordered after ClEp.

Finally, note that when 3 in (502) is /w/, there is a contrast between types a) and b).

---

[38] Also BP /kəmbra:m/ <SP camarão and BP /təmbri:m/ < SP tamarim. The epenthetic /b/ is a quite early development. Cf. Damão and Northern Indo-Portuguese cambrão. Mangalore and Cochin Indo-Portuguese cambrom. Note that BP /ta:mər+ə/ <SP tâmara /támara/ is an exception to this otherwise quite general historical process.

(508) /əwərə/ ——> [ávərə] 'now' Cf. SP agora.

(509) /pəla:wr+ə/ ——> [pəlá:vrə] 'word'

Another epenthesis process is the insertion of [u] at a word-internal boundary to prevent the formation of certain atypical clusters. The phenomenon is not widespread, being found only with the derivational suffixes, agent-forming /+do:r/ and noun-forming /+mæ:nt+u/, and with the locative inflectional suffix, /#ntu/. It is, therefore, open to question whether this process is actually phonologically based or whether some morphological conditioning is also at work. It is certainly possible to formulate general rules which brea certain types of clusters and which will account for the data. The data do not, however, force an unique formulation. Moreover, it is not possible to escape with less than three separate specifications of the environment: one for each of the affixes involved. With these caveats in mind, we tentatively offer the following rule:

(510) Transboundary Epenthesis (TbEp):

$$\emptyset \longrightarrow \begin{bmatrix} V \\ +back \\ +high \end{bmatrix} / \left\{ \begin{array}{ll} \left\{ \begin{bmatrix} -son \\ +voice \end{bmatrix} \right\} \_ \begin{bmatrix} -son \\ +voice \end{bmatrix} & a) \\ \begin{bmatrix} -son \\ -cont \end{bmatrix} \_ [+nas] & b) \\ [+cons]\# \_ [-cont]_2 & c) \end{array} \right.$$

(511) a) /julg+a:/ ——> [ǰulgá·] 'judge' V

b) /julg+do:r/ ——> [ǰulgudó·r] 'judge' N

(512) a) /ku:sp+i/ ——> [kú:spi] 'spittle'

b) /ku:sp+do:r/ ——> [kuspudó˙r] 'spittoon'

(513) a) /pek+a:d+u/ ——> [peká˙du] 'sin' N

b) /pek+do:r/ ——> [pekudó˙r] 'sinner'

(514) a) /ku:lp+ə/ ——> [kú:lpə] 'guilt'

b) /ku:lp+do:r/ ——> [kulpudó˙r] 'guilty'

(515) a) /ke:s+ə/ ——> [ké:sə] 'complaint, lawsuit'

b) /ke:s+do:r/ ——> [kesudó˙r] 'complainer, plaintiff'

(516) a) /inti:nd+ə/ ——> [intí˙ndə] 'understand'

b) /inti:nd+mæ:nt+u/ ——> [intindumǽ˙ntu] 'meaning'

(517) a) /tent+a:/ ——> [tentá˙] 'tempt'

b) /tent+mæ:nt+u/ ——> [tentumǽ˙ntu] 'temptation'

(518) a) /bist+i:/ ——> [bistí˙] 'dress' V

b) /bist+mæ:nt+u/ ——> [bistumǽ˙ntu] 'dressing' (?)

(519) /[[kinta:l]$_N$ntu]$_N$/ ——> [kintá:l̇untu]

(garden LOC) 'in the garden'

(520) /[[blət]$_N$ntu]$_N$/ ——> [blə́tuntu] 'in tin'

(521) /[[tre:z]$_N$ntu]$_N$/ ——>[tré:zuntu] 'in the three'

(522) /[[suw+o:r]$_N$ntu]$_N$/ ——> [suwó˙runtu] 'in sweat'

(523) /[[a:s]$_N$ntu]$_N$/ ——> [á:suntu] 'in steel'

(524) /[[de:ws]$_N$ntu]$_N$/ ——> [dé:wsuntu] 'in/with God'

Cf.

(525) a) /sər+a:/ ——> [sərá˙] 'saw' V

b) /sər+do:r/ ——> [sərdó˙r] 'sawer'

(526) a) /ba:yl+a:/ ——> [bəylá˙] 'dance' V

b) /ba:yl+do:r/ ——> [bəyldó˙r] 'dancer'

(527) a) mi:d+i:/ ——> [midí˙] 'measure' V

b) /mi:d+i:+do:r/ ——> [mididó˙r] 'surveyor'

(528) a) /ruwin+a:/ ——> [ruviná˙] 'ruin' V

b) /ruwin+a:+do:r/ ——> [ruvinədó˙r] 'spoiler'

(529) a) /ka:z+a:/ ——> [kəzá˙] 'marry'

b) /ka:z+mæ:nt+u/ ——> [kæzmǽ˙ntu] 'wedding'

c) /[[ka:z+ə]$_N$ntu]$_N$/ ——> [ká:zəntu] 'at home'

(530) /[[rəy]$_N$ntu]$_N$/ ——> [réyntu] 'in a row'

(531) /[[a:w]$_N$ntu]$_N$/ ——> [á:wntu] 'in water'

Note that TbEp applies after (761) Nasal Elision II, which deletes one of two nasals before a stop.

(532) /[[dəyn]$_N$ntu]$_N$/ ——> [déyntu] 'in an inch'

(533) /[[ma:m]$_N$ntu]$_N$/ ——> [má:ntu] 'in the hand'

4.2 Elision

Word internally, sequences of two or more vowels do not occur. When, as a result of construction, two vowels come together, a glide may be inserted, the first of the vowels may be reduced or elided, or nothing may happen. When the first of the two vowels is a short unstressed back vowel (only /u/, /ə/, and rarely, /o/ occur) it is usually elided, though in careful speech it may remain, in which case it is usually reduced. If the first vowel is non-low and not /ə/, a glide agreeing with it in backness is introduced between it and the following vowel. Note that the domains of the two processes partially overlap; Glide Epenthesis is

obligatory and hence follows Vowel Elision.[39]

(534) Vowel Elision (VE):

allegro speech

optional

$$\begin{bmatrix} V \\ +back \\ -long \\ -stress \end{bmatrix} \longrightarrow \emptyset \; / \; \_\#_{o}V$$

(535) Glide Epenthesis (GEp):

$$\emptyset \longrightarrow \begin{bmatrix} -cons \\ -syll \\ \alpha back \end{bmatrix} \; / \; \begin{bmatrix} V \\ -low \\ \alpha back \\ \alpha round \end{bmatrix} \; \_\#_{o}V$$

(536) /[tə [ənd+a:]$_{V}$]$_{V}$/ ⟶ [təəndá· ~ təndá·]
(PRES go) 'going'

(537) /[[tə [ənd+a:]$_{V}$]$_{V}$[[ɔ:r+ə]$_{N}$s]$_{N}$]$_{NP}$/ ⟶
(PRES go hour PL
[tə$\overset{2}{\text{ə}}$nda $\overset{1}{\text{ɔ}}$:rəs ~ t$\overset{2}{\text{ə}}$nd$\overset{1}{\text{ɔ}}$:rəs] 'when going'

(538) /[lo [ənd+a:]$_{V}$]$_{V}$/ ⟶ [lowəndá· ~ ləndá·]
(FUT go) 'will go'

(539) /[[si:nk+u]$_{A}$[usæ:nt+ə]$_{A}$]$_{A}$/ ⟶
(five hundred)
[s$\overset{2}{\text{ɪ}}$:ŋkuwus$\overset{1}{\text{æ}}$·ntə ~ s$\overset{2}{\text{ɪ}}$:ŋkus$\overset{1}{\text{æ}}$·ntə] 'five hundred'

---

[39]For the sake of clarity Glide Epenthesis is formalized here rather than in Section 5.

(540) /te:#a:w / ——> [teyá˙w]
(tea water) 'tea'

(541) /[[isti]$_N$[[ung]$_A$ə]$_N$]$_S$/ ——> [^{2}ístiyùŋŋə^{1}]
(this one PRO) 'This is one.'

4.3 Vowel quality

The high and mid front vowels /i i: e e:/ are retracted to [ɨ ɨ: ɜ ɜ:] respectively under certain circumstances. all are retracted before clusters of the type rC. Short /i/ may also be optionally retracted before rV with some speakers. In addition, short [e] is in free variation with [ɜ] in environments other than _ rC.

(542) Retraction (Ret):[40]

$$\begin{bmatrix} V \\ -low \\ -back \end{bmatrix} \longrightarrow [3back]$$

$$/ \left\{ \begin{array}{ll} - \begin{bmatrix} C \\ +son \\ -nas \\ -ant \end{bmatrix} C & \text{(a)} \\ \begin{bmatrix} \text{—} \\ -long \\ <+high> \end{bmatrix} \left\langle \begin{bmatrix} C \\ +son \\ -nas \\ -ant \end{bmatrix} V \right\rangle & \text{(b) optional} \end{array} \right.$$

[40] We assume [u] is [1back] and [4back] = [-back]. The rule cannot be further contracted. Clearly however, a single process is involved, which for the short vowels may take place in a less restricted environment than for the long vowels. The fact that Retraction applies optionally in the (b) environment suggests that the process has recently been extended to this environment. This interpretation is supported by the fact that older speakers apply this part of the rule less frequently than younger speakers. There is a problem here for the generative view of historical linguistics, according to which the extension of processes to new environments should correspond formally to rule simplifica-

(543) /we:rd+i/ ——> [vɜ́:rdi] 'green'

(544) /repert+i:/ ——> [repɜrtí˙ ~ rɜpɜrtí˙] 'divide, share'

(545) /kewr+a:/ ——> [kewrá˙ ~ kɜwrá˙] 'break' V

(546) /i:rt+i/ ——> [ɨ́:rti] 'straight'

(547) /dirt+ə/ ——> [dɨ́rtə] 'melt'

(548) /sirw+i:/ ——> [sɨrví˙] 'work' V

(549) /bew+ə/ ——> [bévə ~ bɜ́və] 'drink' V

(550) /sestə+fərə/ ——> [séstəfərə ~ sɜ́stəfərə] 'Friday'

(551) /jinjiwir+i/ ——> [ǰíñǰivɨri ~ ǰíñǰiviri] 'ginger'

(552) /fir+i:/ ——> [fɨrí˙ ~ firí˙] 'cut open'

Cf. (553) /se:r+ə/ ——> [sé:rə] 'wax' N

(554) /pi:ris/ ——> [pí:ris] 'saucer'

Before a word final /r/ these vowels are not retracted; instead they develop a central offglide.

(555) Offgliding (Offg):

$$\emptyset \longrightarrow \begin{bmatrix} -\text{syll} \\ -\text{cons} \\ -\text{high} \end{bmatrix} \; / \; \begin{bmatrix} \text{V} \\ -\text{back} \\ -\text{low} \end{bmatrix} - \begin{bmatrix} \text{C} \\ +\text{son} \\ -\text{nas} \\ -\text{ant} \end{bmatrix}$$

(556) /de:r/ ——> [dé:^{ə}r] 'daily'

(557) /prum+e:r/ ——> [prumé˙^{ə}r] 'first'

(558) /osi:r/ ——> [osí˙^{ə}r] 'he/she HON'

Cf. (559) /suw+o:r/ ——> [suwó˙r] 'sweat' N

In non-initial syllables, unstressed /u/ is unrounded in

---

tion (cf. King pp. 58-59). Here, however, the formal counterpart to generalization is partial rule addition, as a result of which the new rule is more complicated that the old.

casual speech unless the previous consonant is a labial, or the vowel of the preceding syllable is round.

(560) Unrounding (UR):

casual speech

$$\begin{bmatrix} V \\ +back \\ +high \\ -stress \end{bmatrix} \longrightarrow [-round] \;/\; \begin{bmatrix} V \\ -round \end{bmatrix} \begin{bmatrix} C \\ \alpha ant \\ \beta cor \end{bmatrix}_{o} -$$

Condition: if α is + then β is +

(561) /ka:rt+u/ ——> [ká:rtu ~ ká:rtʉ] 'quarter'

(562) /fæ:r+u/ ——> [fǽ:ru ~ fǽ:rʉ] 'iron'

(563) /inti:nd+mæ:nt+u/ ——> [intindumǽ·ntu ~ intindʉmǽ·ntʉ] 'meaning'

(564) /i:ng+u/ ——> [í:ŋgu ~ í:ŋgʉ] 'asafoetida'

Cf.

(565) /ku:rt+u/ ——> [kú:rtu] 'short'

(566) /o:r+u/ ——> [ó:ru] 'gold'

(567) /kərdəmu:ng+u/ ——> [kərdəmú·ŋgu/ 'cardamon'

(568) /tæ:mp+u/ ——> [tǽ:mpu] 'time'

(569) /surka:c+i/ ——> [surká·či] 'squirrel'

In SP the (long) stressed low vowels [ǒ: ǽ:] alternate with unstressed [o e]. [41] This alternation continues in BP and is extended to include the alternation mentioned above (note 21 p. 75) between [á:] and [ə].

---

[41] [u ə~ɨ] in modern continental SP. Other paradigmatic alternations in which these vowels take part have been lost in BP along with the paradigms. Some vestiges remain: e.g. the pair /nɔ:yw+ə/ 'bride', /no:yw+u/ 'groom'.

(570) a) [pǽ:zu] 'heavy; pound'

b) [pezá˙] 'weigh'

(571) a) [fǽ:ru] 'iron'

b) [feré˙ru] 'blacksmith'

(572) a) [sǽ:ku] 'dry' A

b) [seká˙] 'dry' V

(573) a) [duwǽ˙nsə] 'sickness'

b) [duwensá˙du] 'sick person'

(574) a) [pǽ:dərə] 'stone'

b) [pedriyá˙du] 'ornamented with stones'

(575) a) [lɔ́:ñǰi] 'far'

b) [loñǰú˙rə] 'distance'

(576) a) [dɔ́:wrə] 'fold' N

b) [dowrá˙] 'fold' V

(577) a) [rɔ́:də] 'wheel'

b) [rodiyá˙] 'go around'

(578) a) [nɔ́:mi] 'name' N

b) [nominá˙] 'nominate'

(579) a) [kɔ́:ntə] 'amount'

b) [kontá˙] 'count'

(580) a) [ɔ́:brə] 'profession'

b) [obré˙ru] 'manual worker'

(581) a) [má:ŋgə] 'mango'

b) [məŋgé˙rə] 'mango tree'

(582) a) [ká:zə] 'house'

b) [kəzá˙] 'marry'

(583) a) [rá:lu] 'scraper'

b) [rəlá•] 'scrape'

(584) a) [gá:stu] 'expense'

b) [gəstá•] 'spend; waste'

(585) a) [bá:rvə] 'beard'

b) [bərvé•ru] 'barber'

(586) a) [á:ltu] 'high'

b) [əltú•rə] 'height'

(587) a) [bá:ylu] 'dance' N

b) [bəyldó•r] 'dancer'

(588) a) [ərnəgá•] 'curse, hate'

b) [ərnəgəsá•m] 'cursing'

Although this list is not exhaustive, probably no more than about twenty-five examples of each of the three types of alternation could be found.[42]

It must be the long low vowel of each pair which is underlying, since the alternations [é: ~ e] and [ó: ~ o] also exist courtesy of the rule of Length Loss (431).

(589) a) /ce:r+u/ ——> [čé:ru] 'fragance, (pleasant) smell'

b) /ce:r+a:/ ——> [čerá•] 'smell (pleasantly)'

(590) a) /o:y/ ——> [ó:y] 'eye' N

b) /o:y+a:/ ——> [oyá•] 'see'

Length Loss will also account for the quantity alternations

[42]Counting as one the many examples of verb in [a:] alternating with verb stem in [ə] before a suffix with a long vowel, as in (588). Even so, the frequency of the [a: ~ ə] alternation may exceed our estimate.

in (570) - (588). The quality alternations can be accounted for by a later rule of Raising, which will be formulated shortly. First, however, note that on the phonetic level, unstressed short low vowels do not occur except when stressed /æ:/ or /ɔ:/ follows in the next syllable.[43] Thus as a first approximation, we present the quite general rule:

(591) Raising (Rais): first approximation

$$\left[\begin{smallmatrix} V \\ -\text{stress} \end{smallmatrix}\right] \longrightarrow [-\text{low}] \;/\; [X \_ Y]_L$$

Condition: X and Y contain no word boundary.

Raising is a word level phenomenon and thus does not apply in larger constructions after the application of higher level stress rules. In (592), for example, the low vowel in /ɔ:r+ə/ will be stressed at the word level by MSR and will not undergo raising. Though it is destressed by later rules (see note 32 p. 88), it remains low. It does, however, lose its length via LL, a rule which applies both below and above the word level. For this reason Raising and Length Loss cannot be combined.

(592) /[[[do:z]$_A$[[ɔ:r+ə]$_N$pə]$_N$]$_{NP}$[ka:rt+u]$_N$]$_{NP}$/ ——>

(two hour DAT quarter)

[dɑ̆^{2}:z ɔrəpə kă^1:rtu] 'quarter to two'

The condition on X and Y in (591) is necessary to prevent

[43] For the apparent exception [væstumǽ·ntu] see discussion following (603).

the rule from applying in compounds:

(593) /bɔ:wər # d # a:w / ——> [bɔvərdá˙w]
(pumpkin of water 'watermelon'

The raising of /a:/ in (594) and (595) is accomplished by the phrase-level rule ARed (574).

(594) /pa:n+u#wæ:y/ ——> [pənuvǽ˙y]
(cloth old) 'rag'

(595) /[ja:[wi:]$_V$]$_V$/ ——> [javí˙ ~ jəví˙]
PAS come 'came'

A class of exceptions to Raising as formulated in (591) arises when the vowel to be raised is followed by stressed [ǽ˙] ([ɔ́˙] does not occur) in the next syllable;[44] i.e. in just those positions where unstressed short low vowels are permitted (cf. note 16 p. 59). When the vowel in such an environment is /æ:/ or /ɔ:/, no quality change takes place (though LL still applies).

(596) a) /sæ:t+i/ ——> [sǽ:ti] 'seven'
b) /sæ:t+æ:nt+ə/ ——> [sætǽ˙ntə] 'seventy'
(597) a) /nɔ:w+i/ ——> [nɔ́:vi] 'nine'
b) /nɔ:w+æ:nt+ə/ ——> [nɔvǽ˙ntə] 'ninety'
(598) a) /wæ:nt+u/ ——> [vǽ:ntu] 'wind'
b) /wæ:nt+æ:z+ə] ——> [væntǽ˙zə] 'flatulence'

[44] No examples have come to light of non-initial syllable underlying low vowels in this environment.

(599) a) /pɔ:wr+i/ ——> [pɔ́:wri] 'poor'

b) /pɔ:wr+æ:z+ə/ ——> [pɔwrǽ˙zə] 'poverty'

An /a:/ followed by stressed [ǽ˙] also does not undergo raising; however it does become assimilated to this vowel. Length Loss applies as usual, and the [a: ~ æ] alternations of (600) - (602) result.

(600) a) /ka:z+ə/ ——> [ká:zə] 'house'

b) /ka:z+a:/ ——> [kəzá˙] 'marry'

c) /ka:z+mæ:nt+u/ ——> [kæzmǽ˙ntu] 'marriage'

(601) a) /fra:k+u/ ——> [frá:ku] 'weak'

b) /fra:k+æ:z+ə/ ——> [frækǽ˙zə] 'weakness'

(602) a) /ga:ñ+u/ ——> [gá:ñu] 'profit; gain; interest'

b) /ga:ñ+æ:z+ə/ ——> [gæñǽ˙zə] 'stinginess'

Before examining this assimilation further, we shall reformulate Raising so that it applies only when a non-low vowel or /a:/ follows.

(603) Raising (Rais): final version

$$\begin{bmatrix} V \\ -\text{stress} \end{bmatrix} \longrightarrow [-\text{low}] \;/\; [X\_C_o \begin{bmatrix} \left\{ \begin{matrix} V \\ [-\text{low} \\ [+\text{back} \\ [-\text{round} \end{matrix} \right\} \end{bmatrix} Y]_L$$

Condition: X and Y contain no #

Note that only the vowel of the immediately following syllable is relevant to the rule: Raising takes place as usual in (604) b) despite the presence of [ǽ˙] at two

syllables' distance.

(604) a) /bæ:wd+u/ ——> [bǽ:wdu] 'drunk'

b) /bæ:wd+i:s+mæ:nt+u/ ——> [bɜwdismǽ˙ntu] 'drunkenness'[45]

The troublesome form (605) may be handled as a regular derivation from /wæst+mæ:nt+u/, if we order Raising before Trans-boundary Epenthesis, allowing the former to be blocked by the low stressed [ǽ˙] before epenthetic [u] creates a new second syllable.

(605) [væstumǽ˙ntu] 'windy weather'

There is, however, no independent evidence for such an order, and as (605) is non-alternating, the exclusion of /u/ in the underlying representation is ad hoc. For those preferring a less arbitrary solution a possible alternative would be to mark underlying /wæstu+mæ:nt+u/ as an exception to Raising.

Returning to the assimilation observed in (600) - (602) above, we note that it is a synchronic result of a more widespread phenomenon of vowel harmony, whereby historical unstressed initial syllable a has become assimilated to a

---

[45](604) b) could also be derived from the verb /bew+ə/ 'drink', which already has a mid vowel at the underlying level (see p. 114). Arguing against this are the /d/ also present in the adjective and the fact that the abstract noun-forming suffix /-i:s+ə/ is usually added to an adjectival base.

following stressed *ǽ* or *ɔ́*.[46]

(606) /pænæ:l+ə/ ——> [pænǽ˙lə] 'pot'; cf. SP panela /panɛla/ id.

(607) /æzæ:d+u/ ——> [æzǽ˙du] 'sour'; cf. SP azedo /azedu/ id.[47]

(608) /kɔtɔ:lik+u/ ——> [kɔtɔ́˙liku] 'Catholic'; cf. SP católico id.

(609) /sɔwɔ:l+ə/ ——> [sɔvɔ́˙lə] 'onion'; cf. SP cebola /sebola/ id.[48]

Note that non-initial syllable *a* as well as initial syllable *a* separated from the conditioning environment by one or more syllables were unaffected.

(610) /ənənɔ:s/ ——> [ənənɔ́˙s] 'pineapple'

(611) /əwiyə+mæ:nt+u/ ——> [əviyəmǽ˙ntu] 'tool'

The forms (606) - (609) are non-alternating and thus the

---

[46] Two exceptions are:
/əljɔ:fr+i/ ——> [əlǰɔ́˙fri] 'pearl'; cf. SP aljofre /aljofre/ 'seed pearl'
/ənjɔ:l+ə/ ——> [əñǰɔ́˙lə] 'fishhook'; cf. SP anzol /anzɔl/ id.

[47] The lowering of stressed /e/ and /o/, as in (607) and (609) is a very common phenomenon which is also found in other Portuguese-based creoles, as well as in continental and Brazilian non-standard dialects. Although exceptions are numerous, the general pattern in Sri Lanka Portuguese is (informally):

SP é > BP æ: / _C_o {u, a}

SP ó > BP ɔ: / _C_o {i, a}

[48] We assume intermediate **sabɔla*; cf. CVC [səbɔ́lə] (Principense /sɛbɔ́la/) and modern continental (but not Brazilian) Ptg. [səbɔ́lə].

assimilated vowels appear in the underlying representations. A synchronic vowel harmony rule can be posited only for alternating forms such as (600)- (602) above and (613) below and also in compounds.

(612) Vowel Harmony (VH):

$$\begin{bmatrix} V \\ -long \\ +back \\ -round \end{bmatrix} \longrightarrow \begin{bmatrix} +low \\ \alpha back \\ \alpha round \end{bmatrix} / \# C_o - C_o(\#)C_o \begin{bmatrix} V \\ +low \\ \alpha back \\ \alpha round \end{bmatrix}$$

(613) a) /əkɔ:rd+ə/ ——> [ɔkɔ́ˑrdə] 'agreement'

b) /əkɔ:rd+a:/ ——> [əkordáˑ] 'agree'

(614) a) /ək #ɔ:ndə/ ——> [ɔkɔ́ˑndə]

(that way) 'over there'

b) /ək(+ə)l+ɔ:r+ə/ ——> [əkəlɔ́ˑrə ~ ɔklɔ́ˑrə]

(that hour) 'then'

(615) /pa:s#tæ:mp+u/ ——> [pæstǽˑmpu]

(pass time) 'amusement, party'

Vowel Harmony must be ordered after Raising since underlying /ə/ in (613) b) does not assimilate to underlying /ɔ:/. The two items pointed out in note 46 p. 112 will be marked as exceptions to VH.

There is evidence for an historical process similar to Vowel Harmony which lowered mid /e o/ in the VH environment, but left unchanged their backness and roundness.

(616) /wærgɔ:ñ+ə/ ——> [værgó˙ñə] 'shame, shyness';
cf. SP vergonha /vergoña/ 'shame'

(617) /sægræ:d+i/ ——> [sægrǽ˙di] 'private';
cf. SP segredo /segredu/ 'secret'

Additional examples are the synchronic low vowels in (596) b) - (599) b). The process was not general, however, and has no synchronic analog. The mid vowels in the following did not become lowered.

(618) a) /se:s+æ:nt+ə/ ——> [sesǽ˙ntə] 'sixty'; cf. (596).
b) /o:yt+æ:nt+ə/ ——> [oytǽ˙ntə] 'eighty'

(619) /repɔ:st+ə/ ——> [rʒpó˙stə] 'answer' N

Before leaving the subject of Raising, we should point out that there is a historically related set of alternations exhibited in (620) - (622), between the stressed long low vowels in nouns and adjectives, and their stressed short mid counterparts in corresponding -ə verbs.

(620) a) [prǽ:ndu] 'education'
b) [prέndə] 'learn'

(621) a) [bǽ:ɯdu] 'drunk'
b) [bʒ́və] 'drink' V

(622) a) [tó:si] 'cough' N
b) [tósə] 'cough' V

There are very few, certainly less than a dozen, such related pairs. From a synchronic point of view these alternations are morphologically conditioned, and thus beyond the scope of this work.

5 Consonants

Glide Epenthesis has already been treated in the previous section (rule (535)) and need not be repeated here. Most of the rules diccussed below are low-level phonetic processes, many of which involve optional neutralizations of contrast.

5.1 Strengthening and weakening phenomena

Consonants of the voiceless series are regularly unaspirated and have variants ranging from lax to tense and geminate.[49] Both phonological environment and performance factors are involved in this variation. Generally, in careful speech all voiceless consonants are tense. Intervocalically and as the last element of intervocalic clusters with /y w r/ they tend to be geminate unless the consonant before the preceeding vowel is itself geminate.[50] In rapid casual

[49] Gemination will be interpreted here as length.

[50] Cf. Vasconcellos (1901):

> There is a certain tendency to double an intervocalic voiceless consonant, for example in çapatto, ácca,... chappé. (p. 146.)

Also cf. Dalgado (1900):

> SS representative of x is normally perceived in pronunciation: cassão (pron. cas-são) = caixao, basso (pron. bas-so) = abaixo, pussá (pron. pus-sá) = puxar. Also fessão (pron. fes-são) = feição. Compensation for the simplified diphthong. ...

speech, non phrase initial voiceless consonants tend to be lax. In the environment mentioned above, however, they may be tense or tense and geminate, especially when serving as a kind of point d'appui for an adjacent stressed vowel: a voiceless consonant may even geminate under the influence of stress on a previous vowel at several syllables' distance, provided no obstruent, vowel or lateral intervenes (cf. (640) below). In all speaking styles the tendency to geminate is especially strong after short stressed vowels. It is reasonable to assume that underlying voiceless consonants are [+tense, -long]. The low-level rules (623) and (624) are an attempt to capture the distribution just described.

(623) Laxing (Lax):

allegro speech

optional

$$\begin{bmatrix} C \\ -\text{voice} \end{bmatrix} \longrightarrow [-\text{tense}]$$

Note: especially common / $\begin{bmatrix} +\text{cons} \\ \{+\text{nas}, +\text{lat}, -\text{son}\} \end{bmatrix}$ [-stress] __ [-stress]

TT intensifies the sound, as in Italian, dettá (pron. det-tá) = deitar, battá (pron. bat-tá) = bater. (pp. 22-3.)

From Dalgado's transcriptions such as súppodu (p. 18) and soffrá (p. 18) etc. we can deduce that gemination was not limited to /s t/.

(624) Gemination (Gem):[51]

optional

$$\begin{bmatrix} C \\ +\text{tense} \end{bmatrix} \longrightarrow [+\text{long}] \; / \; [-\text{long}] \begin{bmatrix} V \\ \alpha\text{long} \\ \text{mstress} \end{bmatrix} \begin{bmatrix} +\text{son} \\ -\text{nas} \\ -\text{lat} \end{bmatrix}_0 - \begin{bmatrix} V \\ \text{nstress} \end{bmatrix}$$

Note: especially common if α is - or m or n stronger than 0.

(625) /sa:k+u/ ——> [sá:kku ~ sá:ku] 'bag'

(626) /fe:c+i/ ——> [fé:čči ~ fé:či] 'lock' N

(627) /sæ:t+i/ ——> [sǽ:tti ~ sǽ:ti] 'seven'

(628) /cəp+ə/ ——> [čə́ppə ~ čə́pə] 'hat'

(629) /po:rk+u/ ——> [pó:rkku ~ pó:rku] 'pig'

(630) /o:yt+u/ ——> [ó:yttu ~ ó:ytu] 'eight'

(631) /surka:c+i/ ——> [surkká•či ~ surká•čči] 'squirrel'

(632) /[illegible]:t+u/ ——> [səppá•tu ~ səpá•ttu ~ səpá•tu] 'shoe'

(633) /[illegible]+i/ ——> [dó:ssi ~ dó:si] 'sweet'

(634) /[illegible]ti:z+m+u/ ——> [bəwttí•zmu ~ bəwtí•zmu] 'baptism'

(635) /gəwrt+a:/ ——> [gəwrttá• ~ gəwrtá•] 'dig, investigate'

(636) /bɔ:f+i/ ——> [bó:ffi ~ bó:fi] 'lungs'

(637) /burf+a:/ ——> [burffá• ~ burfá•] 'drizzle, sprinkle'

---

[51] As the geminate consonant which inhibits (624) can only have arisen via the same rule, we must assume a left to right application of the rule. This would be predicted by Howard's directional theory of rule application (cf. Howard 1975, p. 111) since the relevant environment is to the left of the segment undergoing the rule.

The following are taken from recorded natural conversations:

(638) /sæ:t+æ:nt+ə#si:nk+u/ ——> [sætt$\overset{3}{\underset{\circ}{\text{æ}}}$ndə$\overset{1}{\underset{\circ}{\text{z}}}$i:ŋ$\underset{\circ}{\text{g}}$u]
(seventy five) 'seventy-five'

(639) /[ki:]$_{NP}$[po:y]$_{AUX}$[fəyə]$_{V}$/ ——> [$\overset{2}{\text{k}}$i: ppoy $\overset{1}{\text{f}}$əyə]
(what can do 'What can we do?'

(640) /[[[əwərə]$_{ADV}$[tə[ma:c+a:]$_{V}$]$_{V}$]$_{S}$[mɔ:stər]$_{N}$]$_{NP}$/ ——>
(now PRES go way)
[$\overset{2}{\text{ə}}$vərə ttə$\overset{3}{\text{m}}$ə$\overset{1}{\text{č}}$a mɔ:stər] 'the way [things] are going now'

(641) /[[o:y]$_{N}$[tä:m]$_{C}$]$_{NP}$[[[ni+kər+ə]$_{AUX}$[pərs+ə]$_{V}$]$_{S}$[ley]$_{C}$]$_{ADV}$
(eye also won't see like
[te:m]$_{V}$/
is)
——> [oy ttä•n nikərə $\overset{2}{\underset{\circ}{\text{b}}}$ərssə ley $\overset{1}{\underset{\circ}{\text{d}}}$ẽ:]
'My eye also doesn't seem to see properly.'

The nasals also optionally occur geminate intervocalically, but only after short stressed vowels in initial syllables. The stronger the stress, the greater the tendency to geminate. We shall assume that underlying nasals are redundantly [-tense]. Since geminate nasals occur in a strict subset of the environments in which geminate voiceless consonants occur, we may posit a rule which operates after stress assignment to make nasals tense in geminable

position and then allow a subsequent application of Gemination to make them optionally [+long]. This two-stage analysis is supported by the fact that Palatal Weakening (rule (678) below), which applies to [-tense] palatals, does not apply to /ñ̃/ in geminable position, even when this segment optionally does not undergo Gemination.

(642) Nasal Tensing (NT):

$$[+\text{nas}] \longrightarrow [+\text{tense}] / \#\#C_0 \begin{bmatrix} V \\ -\text{long} \\ +\text{stress} \end{bmatrix} \_ \ (\#)V$$

The optional word boundary allows tensing to take place before a vowel-initial suffix, as in (644) and (646). In the latter (642) applies to [ŋ] formed from /ng/ by (730) and (747).

(643) /kum+ə/ ——> [kúmmə ~ kúmə] 'eat'

(644) /[[məm]$_N$a:]$_N$/ ——> [mə́mma ~ mə́ma] 'mother'

(645) /tiñ+ə/ ——> [tíñ̃ə ~ tíñə] 'was'

(646) /[[ung]$_A$ə]$_{NP}$/ ——> [úŋŋə ~ úŋə] 'one'

Cf.

(647) /kum+e:r+ə/ ——> [kumé'rə] 'food'

(648) /[nun[tiñə]$_V$]$_V$/ ——> [nuntíñə ~ nuntíỹə] 'wasn't'

(649) /u:ñ+ə/ ——> [ú:ñə] '(finger) nail'

As a consequence of Nasal Tensing and Gemination, geminate consonants will not appear in underlying representa-

tions, except in (650).[52]

(650) /buyy+u/ ——> [búyyù] 'monkey'

The articulation of the voiced lax obstruents [b d g] is very weak intervocalically and in intervocalic clusters with r. In this environment spirantization often occurs. The segment /j/ does not participate in this process but may undergo a different weakening process which will be discussed below.

(651) Spirantization (Spir):

optional

$$\begin{bmatrix} \alpha\text{high} \\ \alpha\text{back} \\ -\text{tense} \\ -\text{nasal} \end{bmatrix} \longrightarrow [+\text{cont}] \;/\; V \left(\begin{bmatrix} +\text{son} \\ -\text{nas} \\ -\text{lat} \end{bmatrix}\right)\_\_ \left(\begin{bmatrix} +\text{son} \\ -\text{nas} \\ -\text{lat} \end{bmatrix}\right) V$$

The non-coronal stops /b/ and /g/ become spirantized much more frequently than does coronal /d/. Spirantization is more frequent in the speech of younger speakers than in that of older speakers. Not reflected in the rule is the fact that the spirant viariants of /b/ and /g/ need not be strictly homorganic: /b/ has variants [β~v][53] and /g/ has the variants [ɣ~x~h]. The further rules necessary to make

---

[52] It is possible that this form too could be accounted for by an extension of the domain of Nasal Tensing, but I do not have the data to decide whether such is a valid solution. In any case the rule could no be extended to cover all sonorants, since we have [éli] 'he' and [élə] 'she'.

[53] I have no examples of free variation before o, e or u; usually only [b] is found in the environment _Vw. See the discussion p. 131 for a further statement on the overlap of /b/ and /w/.

the required adjustments will be dispensed with here. In the examples below, only actually observed phonetic shapes are given; the other variants are presumed to be possible, but lacking in the corpus. Spirantization must be ordered before Laxing to prevent its applying to laxed voiceless stops.

(652) /sub+i:/ ——> [subí· ~ suví·] 'rise'
(653) /ka:br+ə/ ——> [ká:bərə ~ ká:βərə ~ ká:vərə] 'goat'
(654) /sɔ:gr+ə/ ——> [sɔ́:grə ~ sɔ́:ɣərə] 'mother-in-law'
(655) /sægræ:d+i/ ——> [sæɣrǽ·di] 'private(ly)'
(656) /iya:rgə/ ——> [iyá·rgə ~ iyá·rɣə] 'beside'
(657) /ka:g+a:/ ——> [kəgá· ~ kəɣá· ~ kəhá·] 'crap' V
(658) /sæ:g+u/ ——> [sǽ:gu ~ sǽ:ɣu ~ sǽ:xu] 'blind'
(659) /ceg+a:/ ——> [čegá· ~ čehá·] 'arrive'
(660) /pa:g+ə/ ——> [pá:gə ~ páɣə ~ pá:hə] 'pay' N
(661) /po:dr+i/ ——> [pɔ́:dri ~ pɔ́:ðri ~ pɔ́:ðiri] 'rotten'

In many items [b] varies freely with [v] in initial position before a vowel. Other items have only [v]; still others only [b]. Which of the three possibilities will occur cannot generally be predicted. As some of the invariant forms are very common items, it is unlikely that the corpus is deficient. It will therefore be necessary to represent the variations lexically.

(662) /$^{b}/_{w}$o:s/ ——> [bó:s ~ vó:s] 'you [-HON, -PL]'

(663) /bo:m/ ——> [bó:m] 'good'

(664) /ba:s+u/ ——> [bá:su] 'low'

(665) /wəz+a:/ ——> [vəzá˙] 'pour'

(666) /$^{b}/_{w}$a:k+ə/ ——> [bá:kə ~ vá:kə] 'cow, bull'

(667) /$^{b}/_{w}$ist+i:/ ——> [bistí˙ ~ vistí˙] 'dress' V

(668) /wi:/ ——> [ví:] 'come'

(669) /bi:sp+u/ ——> [bí:spu] 'bishop'

Direct lexical representation will also be used to account for the few instances in which [g] alternates or varies with [w], [y], or Ø:

(670) a) /fu$^{g}/_{y}$+e:r+ə/ ——> [fugé˙rə ~ fuyé˙rə] 'banana tree'

b) /fi:w/ ——> [fí:w] 'banana'

(671) a) /præ:g+a:/ ——> [pregá˙] 'nail' V

b) /præ:w/ ——> [prǽ:w] 'nail' N

(672) /nu+$^{g}/_{w}$ə/ ——> [núgə ~ núwə] 'isn't'

(673) /əgɔ:rə ~ əwərə/ ——> [əgó˙rə ~ ówərə] 'now'

(674) /a:gu ~ a:w/ ——> [á:gu ~ á:w] 'water'

(675) /pe$^{g}/_{y}$+a:/ ——> [pegá˙ ~ peɣá˙] 'catch'

(676) /lər(g)+a:/ ——> [lərgá˙ ~ lərɣá˙ ~ lərxá˙ ~ lərá˙] 'leave'

(677) /pr(ig)i:s+ə/ ——> [priɣí˙sə ~ prí:sə] 'laziness'

The reader may have observed that the number of irregularities could be reduced by a rule which changes /g/ to /w/

word finally. (Non-alternating /g/ does not occur in this environment.) Such a rule would be excess baggage, however, as the only forms exhibiting the alternation are (670), (671), and (674).

The voiced palatals are another class of consonants which undergo weakening. Of the SP palatals, /ñ/ and /č ǰ/ (modern SP /š ž/) have been retained, and /l̃/ has become BP /y/.[54] Like the other voiceless stops BP /c/ (i.e. [č]) is unaspirated; both /c/ and /j/ have very little affrication. Non-geminable /ñ/ (see (642) above) in intervocalic position optionally weakens to [ỹ]; intervocalic and, less frequently, initial /j/ optionally weaken to [y]. As /ñ/ does not occur initially, we may assume that the same process is involved in both cases. The rule given below must follow the tensing of geminable /ñ/ by (642) and precede the laxing of /c/ by (623).

(678) Palatal Weakening (PalWk):

optional

$$\begin{bmatrix} +\text{cor} \\ -\text{ant} \\ -\text{tense} \\ \langle +\text{nas}\rangle \end{bmatrix} \longrightarrow \begin{bmatrix} -\text{cons} \\ \langle +\text{nas}\rangle \end{bmatrix} \; / \; \left\{ \begin{matrix} V \\ \# \end{matrix} \right\} \_ \; V$$

The features [+son, -cor] will be supplied by the SPE universal linking rule (XXXV).[55] The same convention would also add [-nas] giving the output [y] in all cases if (678)

[54] The actual development was l̃ > ǰ > y. For further details see discussion of the palatals in chapter III pp. 205 ff.

[55] SPE p. 407.

did not specify that [+nas] segments retain their nasality.

(679) /ga:l+i:ñ+ə/ ——> [gəlí·ñə ~ gəlí·ỹə] 'hen'

(680) /monta:ñ+ə/ ——> [montá·ñə ~ montá·ỹə] 'mountain'

(681) /læ:ñ+ə/ ——> [lǽ:ñə ~ lǽ:ỹə] 'young coconut'

= BT /eLaNi:/

(682) /diñe:r+u/ ——> [diñé·ru ~ diỹé·ru] 'money'

(683) /gre:j+ə/ ——> [gré:ǰə ~ gré:yə] 'church'

(684) /ju:nt+u/ ——> [ǰú:ntu ~ yú:ntu] 'with, joined'

(685) /[ja:[wi:]$_V$]$_V$/ ——> [ǰaví· ~ yaví·]

PAS come 'came'

(686) /[nun[tiñ+ə]$_V$]$_V$/ ——> [nuntíñə ~ nuntíỹə]

NEG was 'wasn't'

Cf.

(687) /tiñ+ə/ ——> [tíññə ~ tíñə] 'was'

Note that both Spirantization and Palatal Weakening have differing phonological consequences depending on the segment to which they apply. When, for example /g/ spirantizes to [ɣ ~ x ~ h] or /ñ/ weakens to [ỹ], no neutralization occurs. On the other hand, the spirantization of /b/ to [v] and the weakening of /j/ to [y] result in the partial loss of the contrasts /b/ : /w/ and /j/ : /y/.

Intervocalic /b/ and /j/ are the residues of incomplete historical changes; /j/, especially, is rare in this position. Both mergers are spreading to initial position, but both /b/ and /j/ are still stable in the clusters

/mb(r) nj lj/.

5.2 Other rules

Another contrast which appears to be on the way out is that between /s/ and /z/. Phonetically /z/ is lax and optionally voiceless (rule (688) below); /s/ is usually tense and may also be geminate (rule (624) above). The difference is, therefore, often only a matter of tenseness. In rapid casual speech even this minimal contrast may be lost in unstressed environments (rule (623) above).

(688) /z/ Devoicing I (ZsI):

optional

$$\begin{bmatrix}+\text{strid}\\+\text{cont}\end{bmatrix} \longrightarrow [-\text{voice}]$$

(689) /ɔ:nz+i/ ——> [ɔ́:nzi ~ ɔ́:nz̥i] 'eleven'

(690) /dɔ:z+i/ ——> [dɔ́:zi ~ dɔ́:z̥i] 'twelve'

(691) /reza:m/ ——> [rezá•m ~ rez̥á•m] '(common) sense'

Cf.

(692) /duw+æ:ns+ə/ ——> [duwǽ•nsə] 'illness'

(693) /dɔ:s+i/ ——> [dɔ́:ssi ~ dɔ́:si ~ dɔ́:z̥i] 'sweet'

(694) /pesa:m/ ——> [pessá•m ~ pesá•m ~ pez̥á•m] 'person'

The situation is made more complicated by the fact that /s/ and /z/ take part in various voicing assimilation processes. The paucity of examples, however, makes formulation of hard and fast rules risky; we shall, therefore, limit ourselves to a discussion of the data.

In word internal, preconsonantal position there is no contrast between [z] and [s]. As illustrated by (695), /s/ may be voiced or voiceless before a voiced consonant. This suggests that /s/ in this environment /s/ is changed obligatorily to z and subsequently optionally devoiced by the low level /z/ Devoicing rule. All examples involve verbs in /i:s/ before the suffix /+mæ:nt+u/.

(695) a) /gərdi:s+ə/ ——> [gərdí˙sə ~ gərdí˙ssə] 'be grateful'

b) /gərdi:s+mæ:nt+u/ ——> [gərdizmǽ˙ntu ~ gərdiz̥mǽ˙ntu] 'gratitude'

There is some scant evidence for the reverse process - the assimilation in voicing of /z/ to a following voiceless consonant:

(696) a) /diz+owdi:s+ə/ ——> [diz̥owdí˙s(s)ə] 'disobey'

b) /diz+gərdi:s+ə/ ——> [di$^{z}/_{z}$gərdí˙s(s)ə] 'be ungrateful'

c) /diz+ka:ns+a:/ ——> [diskənsá˙] 'relax, rest'

Given only b) and c) we could easily set up underlying /dis+/ as the prefix in (696). The crucial example a) appears in my corpus only twice, both times with voiceless [z̥], which could easily be a misinterpreted [s]. Historical evidence, of course, points to /z/: cf. SP desobedecer /dezobedeserr/ 'disobey', SLCP circa 1823[56]

[56] Callaway

disobdeci´do 'disobedient'. The fact that TbEp does not apply also supports /z/ (though weakly, as TbEp is itself only tentatively formulated); cf. (697).

(697) a) /ke:s+ə/ ——> [ké:s(s)ə] 'complaint, lawsuit'

b) /ke:s+do:r/ ——> [kesudó˙r] 'complainer, plaintiff'

In (697) b) prior application of TbEp bleeds the assimilation of /s/ to [z].

The devoicing of /z/ also regularly takes place across word boundaries before a voiceless consonant or a pause. (Note that word final /z/ occurs only in numerals.)

(698) a) /se:z/ ——> [sé:s] 'six'

b) /se:z+e:r/ ——> [se$^{z}_{\underset{\circ}{}}$é˙^{ə}r] ' sixth'

c) /[se:z]$_A$[ɔ:r+ə]$_N$/ ——> [sé^{2}:$^{z}_{\underset{\circ}{}}$ ɔ^{1}:rə]

(six hour) 'six o'clock'

d) /[se:z]$_A$[pesa:m]$_N$/ ——> [sé^{2}:s pesá^{1}˙m]

(six person) 'six people'

(699) a) /diz#o:yt+u/ ——> [di$^{z}_{\underset{\circ}{}}$ó˙ytu]

(ten eight) 'eighteen'

b) /diz#se:z/ ——> [dissé˙s] 'sixteen'

In (700) the [s] before the suffix /+æ:nt+ə/ cannot be predicted on phonological grounds.

(700) [ses(s)ǽ˙ntə] 'sixty'

The distinction between /s/ and /z/ is maintained before a word boundary when a voiced consonant or a vowel follows.

(701) a) /[ma:s]$_A$[[di:y+ə]$_N$s]$_N$/ ——> [m$\overset{2}{a}$:s d$\overset{1}{i}$:yəs]
(more day PL 'more days'

b) /[tre:z]$_A$[di:yə]$_N$/[57] ——> [tr$\overset{2}{e}$:$^{z}/_{\mathring{z}}$ d$\overset{1}{i}$:yə]
(three day) 'three days'

(702) a) /[[ung]$_A$[me:s]$_N$]$_{NP}$[a:ntrə]$_P$/ ——>
(one month within)
[$\overset{2}{ũ}$m m$\overset{3}{e}$:s $\overset{1}{a}$:ntrə]
'within a month'

b) /[tre:z]$_A$[a:n+u]$_N$/ ——> [tr$\overset{2}{e}$:$^{z}/_{\mathring{z}}$ $\overset{1}{a}$:ṇu]
(three year) 'three years'

Before a word boundary followed by a consonant /s/ and /z/ may sometimes become [h] or be elided. The phenomenon is much more common with /z/ followed by a voiced consonant.

(703) /[tre:z]$_A$[di:y+ə]$_N$/ ——> [tr$\overset{2}{e}$:$^{z}/_{\mathring{z}}$ d$\overset{1}{i}$:yə ~ tr$\overset{2}{e}$: d$\overset{1}{i}$:yə]
(three day) 'three days'

(704) /[nə#ma:s]$_C$[me:]$_C$/ ——> [n$\overset{2}{ə}$mas m$\overset{1}{e}$: ~ n$\overset{2}{ə}$mah m$\overset{1}{e}$:]
(only EMPH) '[That's] all.'

The glide /w/ is usually rounded after /u/ or /o/. In other environments it has a variant (symbolized here by

---

[57] When a numeral is present, the overt plural marker does not occur.

[w]) with minimal, if any, rounding. We shall assume that rules assigning scalar values for [round] would take care of this variation.[58] In pronouncing [w] the lower lip may approach the upper teeth, but not closely enough to produce friction. A variant [v] of /w/, which has light frication does occur in initial position and optionally in medial position intervocalically and in clusters with /y/ or with liquids; in the last case, if the liquid <u>follows</u> /w/ the following vowel must be unstressed. This distribution is formalized by (705). Phonologically /w/ is clearly a glide. Its distribution, though not identical to that of /y/, is much less like that of /f/[59] and further, in its role in rules such as (623) Laxing, (624) Gemination, and (651) spirantization, it forms a natural class with /y/.

(705) Glide Constriction (GlC):

$$\begin{bmatrix} -\text{syll} \\ +\text{son} \\ +\text{round} \end{bmatrix} \longrightarrow \begin{bmatrix} -\text{cons} \\ -\text{son} \\ -\text{back} \\ -\text{high} \\ +\text{ant} \\ +\text{voice} \end{bmatrix}$$

$$/ \begin{cases} \#\_ & \text{(a)} \\ \text{V}([+\text{son}])\ \_\ \left\langle \begin{bmatrix} +\text{son} \\ -\text{nas} \end{bmatrix} \right\rangle \begin{bmatrix} \text{V} \\ \langle -\text{stress} \rangle \end{bmatrix} & \text{(b) opt.} \end{cases}$$

The feature [+voice] must be specified in the output of the

---

[58] Possibly the unrounding of /u/, discussed in Section 4, should be treated in conjunction with this.

[59] E.g. /w/ and /y/ can occur syllable finally and in clusters with almost any other class of consonant: /f/ is restricted to syllable initial position and to clusters with liquids and nasals.

rule in order to block the SPE universal linking rule (or UMC) (XXI)[60] which supplies [-voice] when triggered by the assignment of [-son]. UMC (XXXIX)[61] will be triggered by the assignment of [-high] to supply [-round]. In the environment of (705) (b) the first optional sonorant need not be specified as [-nas] since nasal plus /w/ clusters do not occur. Examples follow.

(706) /wi:/ ——> [ví:] 'come'

(707) /we:rd+i/ ——> [vʒ́:rdi] 'green'

(708) /kərwi:l+ə/ ——> [kərví˙lə ~ kərwí˙lə] 'bitter gourd'

(709) /pilw+a:/ ——> [pilvá˙ ~ pilwá˙] 'beat'

(710) /a:lwr+i/ ——> [á:lvri ~ á:lwri] 'tree, plant'

(711) a) /dɔ:wr+ə/ ——> [dɔ́:vrə ~ dɔ́:wrə] 'fold' N

b) /dɔ:wr+a:/ ——> [dowrá˙] 'fold' V

(712) a) /əwr+i:/ ——> [əwrí˙] 'open' V

b) /[[[tə [əwr+i:]$_V$]$_V$]$_S$[[ɔ:r+ə]$_N$s]$_N$]$_{NP}$/ ——>

(PRES open hour PL

[t^{2}əwri y ^{1}ɔ́:rəs ~ t^{2}əvri y ^{1}ɔ́:rəs]

'when opening'

(713) /sɔwɔ:l+ə/ ——> [sɔvɔ́˙lə ~ sɔwɔ́˙lə] 'onion'

(714) /æwæ:rt+u/ ——> [ævǽ˙rtu ~ æwǽ˙rtu] 'open' Adj

(715) /duw+æ:ns+ə/ ——> [duwǽ˙nsə ~ duvǽ˙nsə] 'sickness'

(716) /məruwi:y+ə/ ——> [məruvi˙yə ~ məruwi˙yə] 'mischievous'

---

[60] P. 406.

[61] P. 407.

(717) /əwn+a:/ ——> [əwná·] 'fan' V

(718) /de:ws/ ——> [dé:ws] 'god'

(719) /mursæ:w/ ——> [mursǽ·w] 'bat'

(720) /gəwrt+a:/ ——> [gəwrtá·] 'dig'

(721) /frwi:t+ə/ ——> [frwí:tə] 'fruit'

(722) /kwæ:ntr+u/ ——> [kwǽ:nturu] 'coriander'

From a comparison of the foregoing and the discussion of the spirantization of /b/ above (pp. 120 ff.). it is clear that in BP /b/ is in the process of falling together with /w/ everywhere.[62] Even in continental Portuguese, the [v]:[b] contrast has not always been a stable one. In BP the confused state of affairs is continued. The situation is slightly different in initial and non-initial positions, but evidence indicates that in both cases the change is moving through the lexicon. The optional spirantization of medial /b/ leads to words with free variants involving [b ~ β ~ v]; [v] is also one of the possible realizations of /w/. This latter also has glide variants [w ~ ẘ] which the former seems to lack: for example, (652) /sub+i:/ is usually realized as [suví·]; *[suwí·] does not occur in the corpus. Thus intervocalic /b/ and /w/ apparently still contrast, however minimally. There are nevertheless many words in which historical /b/ has changed to /w/ in BP:

(723) /sɔwɔ:l+ə/ 'onion' SP cebola

[62] In the cluster /mb/, /b/ is unaffected by Spirantization. However, there is no /mw/ cluster and thus no contrast to maintian.

(724) /bɔ:wr+ə/ 'pumpkin' SP abóbora

In initial position [v] is the only possible realization of /w/ as Glide Constriction is obligatory: there is no regular phonetic process to account for the free variation [b ~ v] found in some words, as Spirantization applies only to medial voiced stops. In fact, not all initial /b/'s alternate with /w/'s. The alternatives must therefore be represented in the lexicon. See (662) - (669) and accompanying discussion.

We shall now turn to a discussion of various phenomena related to nasals. The reader is reminded that nasals were involved in (642) Nasal Tensing (and (624) Gemination) and (678) Palatal Weakening treated above.

Word final /m/ is usually elided, leaving nasalization of the vowel, when the following word begins with a vowel or /y/, and less frequently before a pause (represented by the symbol |). Rule (725) applies after (678) Palatal Weakening and (705) Glide Constriction, but before (747) Final Cluster Simplification.

(725) Nasal Elision I (NElI): optional

SD: $V\begin{bmatrix}+\text{nas}\\-\text{cor}\end{bmatrix} \# \left\{ \begin{matrix} | \\ [-\text{cons}] \end{matrix} \right.$

1 2 3 4

SC: a) 1 ——> [+nas]

b) 2 ——> Ø

(726) /əmiya:m# o:tər#di:y+ə/ ——> [əmiyã otər dí˙yə]
(tomorrow other day) 'the day after tomorrow'

(727) /[[[mĩñ+ə]$_A$]$_{NP}$[tir+a:]$_V$[ta:m]$_{AUX}$]$_S$[ja:[foy$_V$]$_V$]$_V$/ ——>
(mine take MID PAS went)
[mĩññə tira tã ya fõy^{1}] 'Taking <u>mine</u> along he went.'

(728) /[se:m]$_{AUX}$[ənd+a:]$_V$/ ——> [sẽ^{2}: əndã^{1}˙]
(without go) 'without going'

(729) /[ək+ə]$_N$[ta:m]$_C$/ ——> [^{2}kə tã^{1}:]
(that too) 'That too.'

In word final position, the non-coronal nasal /m/ assimilates in point of articulation to a following consonant. /ñ/ does not occur word finally, and as (741) demonstrates, word final /n/ does not undergo assimilation. Word final [ŋ], derived from /ng/, does assimilate as in (755) - (758). Morpheme internally, only assimilated nasal and consonant clusters are found.[63] It will be assumed that the underlying nasal in bilabial clusters is /m/ and in all other clusters /n/.[64]

---

[63] An exception to this statement is the form /kums+a: ~ kuns+a:/, in which some speakers have the cluster [ms] and others [ns].

[64] Including the palatal clusters [ñč ñǰ], since /ñ/ does not otherwise occur syllable finally.

(730) Nasal Assimilation I (NAssI):

$$\begin{bmatrix}+\text{nas}\\+\text{cor}\end{bmatrix} \longrightarrow \begin{bmatrix}\alpha\,\text{ant}\\ \beta\,\text{cor}\\ \gamma\,\text{high}\\ \delta\,\text{dist}\end{bmatrix} \;/\; - \begin{bmatrix}\text{C}\\ \alpha\,\text{ant}\\ \beta\,\text{cor}\\ \gamma\,\text{high}\\ \delta\,\text{dist}\end{bmatrix}$$

(731) Nasal Assimilation II (NAssII):

$$\begin{bmatrix}+\text{nas}\\-\text{cor}\end{bmatrix} \longrightarrow \begin{bmatrix}\alpha\,\text{ant}\\ \beta\,\text{cor}\\ \gamma\,\text{high}\\ \delta\,\text{dist}\end{bmatrix} \;/\; - \;\#_1 \begin{bmatrix}\text{C}\\ \alpha\,\text{ant}\\ \beta\,\text{cor}\\ \gamma\,\text{high}\\ \delta\,\text{dist}\end{bmatrix}$$

(732) /tæ:mp+u/ ——> [tǽ:mpu] 'time'

(733) /lɔ:nj+i/ ——> [lɔ́:ñǰi] 'far'

(734) /sa:ng+i/ ——> [sá:ŋgi] 'blood'

(735) /konfs+a:/ ——> [koŋfsá˙] 'confess'

(736) /[[ma:m]$_N$su]$_N$[sər+o:t+i]$_N$/ ——> [mā^{2}:nsu sərō^{1}˙ti]
(hand GEN saw) 'handsaw'

(737) /[pikini:m]$_A$[ka:z+ə]$_N$/ ——> [pīkiniŋ kā^{1}:zə]
(small house) 'small house'

(738) /[bo:m]$_A$[[je:nt+i]$_N$s]$_N$/ ——> [bō^{2}:ñ ǰē^{1}:ntis]
(good people PL) 'good people'

(739) /[pər[im]$_N$]$_N$[tə [suw+a:]$_V$]$_V$/ ——> [pərīn təsuwā^{1}˙]
(DAT me PRES sweat) 'I am sweating.'

(740) /[reza:m]$_N$[ley$_{[-MSR]}$]$_C$/ ——> [rezā^{1}˙n ley]
(reason like) 'reasonably'

Cf.

(741) /[[kəlku:n]$_N$pə]$_N$/ ——> [kəlkú˙n pə]
(turkey DAT) 'for the turkey'

Referred to in the previous section, but not formalized there, was the low level rule which inserts a postalveolar apical stop [D] between a nasal and r. The process appears to be obligatory word internally, though (743) and (744) are the only examples of the cluster /nr/. Across word boundaries, the process is optional. The rule, given below, must apply after (496) ClEp since the resultant cluster, [Dr], is not split.

(742) D Epenthesis (DEp):[65] <optional>

$$\emptyset \longrightarrow \begin{bmatrix} C \\ -cont \\ +cor \\ -ant \\ -nas \\ -dist \end{bmatrix} \; / \; \begin{bmatrix} C \\ +nas \end{bmatrix} \; \langle \#_1 \rangle \; \_ \; \begin{bmatrix} -syll \\ -cons \\ -ant \end{bmatrix}$$

(743) /jæ:nr+u/ ——> [ǰǽ:NDru] 'son-in-law'

(744) /tæ:nr+u/ ——> [tǽ:NDru] 'young coconut'

(745) /[ma:m]$_N$[ri:wə]$_P$/ ——> [mā^{2}:N rī^{1}:və ~ mā^{2}:N Drī^{1}˙və]
(hand on) 'on the hand'

It may be though that NAssI and NAssII could be combined

---

[65] I have noticed that one speaker, at least, has extended the domain of application of (742) to include the cluster /lr/, which occurs morpheme internally only in the word /a:lwr+i ~ a:lr+i/. Thus the speaker in question had [á:LDri]; I do not know what he did when a word boundary intervened.

to give (746)

(746) Hypothetical Combined Nasal Assimilation Rule:

$$\begin{bmatrix} +\text{nas} \\ <-\text{cor}> \end{bmatrix} \longrightarrow \begin{bmatrix} \alpha\,\text{ant} \\ \beta\,\text{cor} \\ \gamma\,\text{high} \\ \delta\,\text{dist} \end{bmatrix} \;/\; \_\; <\#_1> \begin{bmatrix} \text{C} \\ \alpha\,\text{ant} \\ \beta\,\text{cor} \\ \gamma\,\text{high} \\ \delta\,\text{dist} \end{bmatrix}$$

In order to see why this is not possible, we first need to discuss the formation of [ŋ].

Phonetic [ŋ] is produced by the action of the Nasal Assimilation rules on nasals followed by velar stops. Rule (747) applies after NAssI to delete the conditioning environment before a word boundary. When a vowel initial word or transformationally added suffix follows, intervocalic [ŋ] results.

(747) Final Cluster Simplification (FCS):

$$\begin{bmatrix} \text{C} \\ +\text{voice} \end{bmatrix} \longrightarrow \emptyset \;/\; [+\text{nas}]\; \_\; \#$$

(748) /$[\text{ung}]_A[\text{a:n+u}]_N$/ ——> [²ŭŋŋ¹ă:ṇu]
(one year) 'one year'

(749) /$[\text{əl+ung}]_A[[\text{ɔ:r+ə}]_N\text{s}\ ]_N$/ ——> [²ə̣luŋ ¹ɔ̆:rəs]
(some hour PL) 'sometimes; perhaps'

(750) /$[\text{ist+i}]_N[[\text{ung}]_A\text{ə}]_N$/[66] ——> [²isti y ¹ŭŋŋə]
(this one) 'This is one.'

---

[66] The suffix /ə/ is added to /ung/ whenever the noun it modifies does not directly follow.

Cf.

(751) /ni:ng+ə/ ——> [ní:ŋgə] 'no one'[67]

(752) /kərdəmu:ng+u/ ——> [kərdəmú'ŋgu] 'cardamom'

In (748) and (750), (642) Nasal Tensing and (624) Gemination have applied. This analysis suffers from a major defect: the only items eligible to undergo (747) are /ung/ and its related forms. This is, however, partially the fault of the data as these forms are the only ones containing [ŋ] not followed by a velar stop. Rule (747) cannot be generalized to include all word final nasal plus stop clusters because of the item /rɔnt/ 'round'; it is interesting, though, that final [nt] in English loans has been simplified in BP:

(753) /konwən/ 'convent'

(754) /winsən/ 'Vincent'

We shall now return to our discussion of Nasal Assimilation. Observe that the word final [ŋ] produced by the above rules assimilates to an initial consonant in the following word:

(755) /[ung]$_A$[rupiy+ə]$_N$/ ——> [ũ^{2}N(D)rũ^{1}piyə]

(one rupee) 'one rupee'

---

[67] Although one informant claimed [ni:ŋgə] was a combination of [ni] '(NEG)' and [uŋŋə], it will be treated here as a monomorpheme. There is no corresponding adjectival form *[ni:ŋ].

(756) /[əl+ung]$_A$[[je:nt+i]$_N$s]$_N$/ ⟶ [^{2}ə̣luñ ^{1}ǰe:ntis]

(some people PL) 'some people'

(757) /[ung]$_A$[fa:k+ə]$_N$/ ⟶ [^{2}ũŋ f^{1}a:kə]

(one knife) 'one knife'

(758) /[[ung]$_A$[[di:y+ə]$_N$pə]$_{NP}$

(one knife DAT

[[ung]$_A$[pæ:z+u]$_N$]$_{NP}$[ba:st+ə]$_V$/ ⟶

one pound suffice)

[^{2}ũn d^{1}i:yəpə | ^{2}ũm p^{3}æ:zu b^{1}a:stə]

'For one day, one pound is enough.'

According to the analysis developed here, in the above items both Nasal Assimilation rules must apply, and therefore, the abbreviation (746) is not possible as the angled brackets imply disjunctive ordering. Moreover, FCS must apply after NAssI and before NAssII. Two derivations are given here to illustrate this point.

(759) /[[ung]$_A$[a:n+u]$_N$]$_{NP}$/ (one year)

1	1	MSR (rule (483))
2	1	NSR (rule (460))
ŋ		NAssI
∅		FCS
(ŋ)		NT (rule (642)) and Gem (rule (624))
	ṇ	Other rules
[^{2}ũŋ(ŋ) ^{1}a:ṇu]		'one year'

(760) /[[ung]$_A$[ma:m]$_N$]$_{NP}$/ (one hand)

1 1	MSR
2 1	NSR
ŋ	NAssI
Ø	FCS
m	NAssII
[$\overset{2}{\text{ũm}}$ $\overset{1}{\text{mã:m}}$]	'one hand'

The analysis is far from pleasing. The only reasonable alternative appears to be to set up underlying /ŋ/ in the morpheme [uŋ]. To do so would certainly eliminate the necessity for (747) Final Cluster Simplification, and allow the Nasal Assimilition rules to be combined. One must weigh this advantage against the additional cost of the extra, minimally functional, phonological unit. It seems to this writer that, as a general principle, the latter should be avoided.

When the locative suffix /ntu/ is added to a nasal final noun, the resultant cluster is simplified to /nt/, We shall arbitrarily assume that it is the second nasal which drops out. When the first nasal is /m/, NAssII will apply, but no

ordering relationship obtains between this and the rule below.

(761) Nasal Elision II (NElII).

$$[+nas] \longrightarrow \emptyset \;/\; [+nas]\ \#\ \_\ C$$

(762) /[[kəlku:n]$_N$ntu]$_N$/ ——> [kəlkú·ntu] 'in the turkey'

(763) /[[ma:m]$_N$ntu]$_N$/ ——> [má:ntu] 'in the hand'

Cf.

(764) /[[mæ:l]$_N$ntu]$_N$/ ——> [mǽ:luntu] 'in the honey'

The alveoler sonorants /n/ and /l/ have developed retroflex allophones following o̲, ɔ̲:, ə̲, and a̲:, intervocalically and finally.[68]

(765) Retroflection (Rflx):

$$\begin{bmatrix}+son\\+cor\end{bmatrix} \longrightarrow \begin{bmatrix}-ant\\-dist\end{bmatrix} \;/\; \begin{bmatrix}V\\+back\\\alpha round\\\beta low\end{bmatrix} - \begin{Bmatrix}V\\\#\end{Bmatrix}$$

Condition: if α is + then β is +

(766) /sɔ:l/ ——> [sɔ́:l̥] 'sun'

(767) /mɔ:l+i/ ——> [mɔ́:l̥i] 'soft'

(768) /ma:l/ ——> [má:l̥] 'bad'

(769) /ɔnkəl/ ——> [ɔ́ŋkəl̥] 'uncle'

(770) /nɔ:n+ə/ ——> [nɔ́:n̥ə] 'woman'

---

[68]The distribution of retroflex [ɳ ɭ] may in fact be wider than this. These phonetic variants, which were not properly investigated until after I had left the field, are difficult to distinguish from the alveolar varieties in a live elicitation situation; on tape they become virtually undifferentiable.

(771) /əla:/ ——> [əl̥á˙] 'there'

Cf.

(772) /mæ:l/ ——> [mǽ:l] 'honey'

(773) /kərwi:l+ə/ ——> [kərví˙lə] 'bitter gourd'

(774) /luna:yr+u/ ——> [luná˙yru] 'lightning'

CHAPTER III: INDIGENOUS INFLUENCE IN BATTICALOA PORTUGUESE

1 Introduction

The tendency of languages which remain in contact over a long period of time to develop similar structural traits is the acknowledged, if not well-understood, phenomenon which results in the formation of linguistic areas (or Sprachbünde), such as South Asia and the Balkans. The mechanism by which such convergent development takes place is presumed to be interference in the speech of multilinguals.[1]

Convergence also takes place in pidginization, but here it differs from the above in several respects. First, it proceeds at a much faster pace. Second, the languages in contact do not themselves change, but a new language is formed which contains facets of all of the contact languages. Finally, significant multilingualism would obviate the necessity for pidgin formation, and thus cannot play a significant role in the convergence; however, the mechanism involved may be something quite similar, such as mother-tongue influence in the unsuccessful attempts of one group to speak the other's language.

Is the distinction between these two types of convergence - that which proceeds slowly in the formation of linguistic areas and that which proceeds rapidly in pidginization - merely an historical one, or are there correlates in the linguistic effects of each? My own opinion is that

[1] Weinreich, p.1. The term multilingual includes bilingual as a limiting case.

they are merely opposite ends of a continuum and cannot be distinguished without some kind of external evidence. The question will not be settled in these pages, though it will be shown, for example, that the current situation with respect to stress and vowel length in BP was not a result of pidginization.

As mentioned in Chapter I, the language introduced to Sri Lanka by the Portuguese was undoubtedly a variety of pidgin Portuguese which had been shaped by their contacts with the peoples of the west coast of India. Indeed, Sri Lanka Portuguese shares many structural traits with other varieties of Indo-Portuguese, although some of these similarities may be due to parallel development caused by the influence of traits common to the languages of the South Asian linguistic area. Also, the possibility cannot be ruled out that innovations in Sri Lanka were carried back into the Indian dialects.

In this chapter I shall demonstrate the influence of indigenous languages of Sri Lanka on the development of Sri Lanka Portuguese, and in particular of Batticaloa Portuguese. The strategy will be to compare Batticaloa Portuguese (BP), Batticaloa Tamil (BT), and Standard Portuguese (SP), to isolate various traits of BP which seem to be due to indigenous influence, and to prove that these are not found in other varieties of Portuguese, including Indo-Portuguese, and hence that that their development is peculiar to Sri Lanka.

In many instances, because Tamil and Sinhala themselves are quite similar, it is not possible to decide whether it is the influence of one or the other which is at work. In general, that of Tamil has been stronger: there are several phonological developments in BP which can be attributed only to Tamil influence, and none which are solely the result of Sinhala influence.

Although the main focus of interest of this work is phonology, the effects on BP of contact with the indigenous languages of Sri Lanka pervade the whole of its grammar. Accordingly, before dealing with phonology, I shall point out some non-phonological congruences between BP and BT (and Sinhala) which appear to have resulted from indigenous influence.

The various coastal dialects of Indo-Portuguese on which data is available, and the sources of the data, are given below. In the rest of the chapter, page references will be made without quoting the source when it is clear to which dialect we are referring.

Diu Spoken on Diu island, off the southern shore of the Kathiawar peninsula. A Portuguese possession until 1961. Schuchardt 1883b.

Daman Spoken in Daman (Damão), 100 miles north of Bombay, a Portuguese possession until 1961. Dalgado 1903.

North IP Spoken north of Goa, including the above two regions. Dalgado 1906; data here is from Thana, just north of Bombay.

Goa Spoken in Goa, a Portuguese possession until 1961. Actually an aberrant variety of SP and not a true creole, and therefore will be ignored in our discussion.

Mangalore Mangalore is 160 mi. south of Goa. Schuchardt 1883c.

Mahé and Cannanore Spoken in Mahé and Cannanore, 100 miles south of Mangalore. As data is insufficient, this dialect will be ignored here. Schuchardt 1889.

Cochin Spoken in Cochin (300 miles south of Mangalore). Schuchardt 1883a.

19th century Sri Lanka Creole Portuguese (19thC SLCP) Dalgado 1900. Data is probably from the Colombo area.

It should be noted that all of this data is based on written materials which can be expected to be conservative. Much of the material given in the above articles is obviously standardized, and I have disregarded such data.

The data on BP is of two types: normal colloquial speech and formal, consciously archaizing speech which my older consultant could produce under stimulus. The BT data is in colloquial style.[2]

2 Non-Phonological Aspects

The best way to introduce this section is to compare a BP text with its BT equivalent. The following is an excerpt from an unrehearsed conversation recorded in BP and

[2] BT, being diglossic, exhibits a far greater number of differences between colloquial and formal styles than BP. We would be straying too far afield to go into these here, however.

later transcribed and translated into BT (see Introduction p. 3):[3]

(1) BP osi:r-su fi:yə -pə portuge:s nuku-səwə
NEG

he GEN daughter DAT Portuguese know

NEGPOT
BT avaŕ -ŕa makaL -ukku portyugi:s[4] teŕi-ya:

A 'His daughter doesn't know Portuguese.'

(2) a) BP nosə ka:zə-ntu se:m pəpiya:
[GEN] without

our house LOC speak

GEN without
BT enka-Ta vi:TT-ila katay-kk-a:ma

BP te:m ma:y -pa:y !?
[PRES]

be mother father CRQ

PRES CNC
BT iŕu-kk-r -a:nka ta:y -takappan !?

B 'In our house are [her] parents not speaking it!?'

---

[3]When giving examples in this section, we shall follow the usual practice of providing both morph-by-morph (occasionally word-by-word) glosses and idiomatic English translations. For examples containing equivalent BP and BT sentences, common interlinear glosses will be given wherever the structure of the two languages permits; structural differences in the two versions will be indicated by glosses which are displaced above or below the line of the common glosses towards the language in question. The identity of the speakers in the excerpts from conversations will be indicated to the left of the idiomatic English translations.

[4]This English loan was preferred by my Tamil consultant to the pejorative paranki [pa:ca] supplied by my BP consultant.

[5]For the sake of brevity, material is omitted between this and the following utterance.

b) BP tə- pəpiya: me: -na:
PRES speak EMPH TAG
PRES CNC
BT katay-kk-r -a:nka ta:n-e:

B 'They are speaking [it], aren't they?'

c) BP əkə poḍiya:s -pə mule:rə
[PL]
those children DAT brain
PL
BT ante piLLay-kaL-ukku mu:La

BP nikara sirwi:
NEGDSC work
work do PRES DESC NEG
BT ve:la cey-r -a -illa

B 'Those childrens' brains don't work.'

d) BP isti nuwə ju:stu isti me: ju:stu fəla:
right
this isn't this EMPH right QUOT
right
BT itu cari illa itu ta:n cari eNTu

B [They say,] 'This isn't right. This is what is right.'

(3) BP iskɔ:lə əndə: -sərə iskɔlə-ntu tudus
school go COND school LOC everyone
DAT PAS
BT paLLi -kku po:-n -a: paLLi -:la ella:rum

BP məlwa:r
Tamil
BT tamiL

C 'If [they] go to school, in school everyone is Tamil.'

(4) a) BP əla: ənda: se: ta:m

there go COND even

PAS

BT anka po:-n -a:l -um

B 'Even if [they] go there,'

b) BP ung luwa:rə |

a place

DAT

BT oŕu iTatt-ukku |

B 'To some place.'|

c) BP tə- junta: luwa:rə ənda: -tu

PRES

meet place go PFC

PRES DESC PAS

BT ku:T-r -a iTatt-ukku po:-y -ŕŕu

B 'Having gone to the place where they meet,'

d) BP nosə je:ntis do:s pesa:m tiñə se:

[GEN] PAS

our people two person be COND

GEN PAS

BT enka-Ta a:kkaL ŕeNTu pe:ŕ iŕu-nt -a:

B 'if two of our people are [there],'

e) BP ja:-pəpiya: se:

PAS

speak COND

PAS

BT katay-cc -a:

B 'if [they] speak [creole],'

f) BP isti mælwa:r-s gra:sə tə- fəyə
PRES
these Tamil PL joke do
PRES CNC
BT inta tamiL -a:ŕ pakuTi paN-r -a:nka

B 'these Tamils make fun [of them]'

g) BP nosə[6] lingwa:y tə- pəpiya: na:
our PRES
language speak TAG
you GEN PRES CNC
BT unka-Ta pa:ca kataɣ-kk-r -i:nkaL-e:

BP əwərə

now

BT ippa

B '[saying], "You are speaking your language, aren't you. Now."'

h) BP enna botus ki: no:s -pə toka:ndu
what (BT) you what us ACC about
BT ni:nka enna enkaL-a paŕŕi

[6]Performance error for /botus-su/ (you-GEN) 'your'.

BP ki: -wo: tə- pəpiya: tə- tira:
PRES PRES take

what INDEF speak

PRES CNC
BT ennam-o: katay-kk-r -i:nkaL ----------[7]

BP fala:

QUOT

BT eNTu

B 'What! you are saying something or other about us (BP: and preparing a trick).'

i) BP əkə wi:də etus wærgɔ:ñə-ntu

that because they shyness LOC

DAT for
BT atu-kk a:ka ve:NTi avanka vekkatt -ila

BP na:- pəpiya:
NEGPOT

speak

INF VOLNEG CNC
BT katay-kk-a ma:TT -a:nka

B 'Therefore, in shyness they won't talk.'

j) BP əkə me: əkə -su ka:ws

that EMPH that GEN affair

BT atu ta:n atu -ŕa vicayam

B 'That's the heart of it.'

---

[7] The equivalent of BP tə-tira: was omitted from the BT translation.

k) BP iskɔ:lə tə- ənda: tæ:mpu -ntu
PRES

school go time LOC

DAT PRES DESC

BT paLLi -kku po:-r -a ka:latt-ila

BP əssi: me: -na:

thus EMPH TAG

BT appaTi ta:n -e:

B 'During school days [lit: in the school-going time] it's like that, of course.'

I do not wish to claim that BP is BT with Portuguese words. This is far from true. Even so, from the above it should be clear that BP and BT show a high degree of structural similarity. It will not be possible in these few pages to investigate thoroughly all the points of congruence. But we shall point out those areas in which the case for indigenous influence is strongest.

2.1 Syntax

In this section we shall investigate word order and some embedding strategies. In both of these areas BP closely parallels BT while SP and other varieties of Indo-Portuguese are usually at variance.

2.1.1 MAJOR CONSTITUENT ORDER. In non-copulative[8] sentences

[8]Copulative sentences, as in other varieties of IP, are often verbless in both languages (e.g. (4), (5), (8), (9)). Some form of the copula is obligatory when, as in (2) d), the sentence is not present and non-negative. The copula is optional in BP in all present non-negative sentences, and in BT in present non-negative sentences not of the NP NP

word order in BT and BP is SOV (e.g. (4) f) h)), except when there is postscripting, contrastive emphasis or extraposition of a 'heavy' embedded clause. Any element may be tacked onto the end of the sentence as a postscript, usually for the purpose of clarification (e.g. (4) g)). Both BP and BT allow free deletion of any non-verbal elements already in the discourse context; these are often re-introduced in a postscript. It is possible to argue that the postscript is not an integral part of the sentence. This view is supported by the intonational evidence: in postscripted sentences as in non-postscripted sentences the main sentence stress and accompanying pitch drop falls on the verb; the postscript forms an unstressed low-pitch coda. In contrastive emphatic sentences such as (2) a), on the other had, the focussed element which follows the verb is best seen as having been extraposed from a preverbal position: here the main sentence stress and accompanying pitch drop fall on the focussed element and the verb is usually unstressed. Finally, heavy embedded clauses are also usually extraposed as in (17) - (18) below.

In SP the unmarked word order is SVO. In other Indo-Portuguese dialects the usual order is also SVO, though most dialects also permit SOV; in addition, Diu and Daman sometimes allow VOS. In 19th century SLCP literature too,

---

type. In BT the copula, like other verbs, is sentence final, but in BP it is optionally non-final. In SP the copula is always olbigatory and is not sentence final.

SVO predominates, but SOV, and occasionally OSV, are also found. In archaizing BP, SVO is common. The data indicate that early IP was mainly SVO, although SOV, the usual South Asian order, may also have been possible. In SLCP, as a result of indigenous influence, the latter became more common, until finally in BP it is the unmarked order.

2.1.2 PP ORDER. Postpositions are characteristic of BP and BT. SP has prepositions exclusively. Other varieties of Indo-Portuguese have mainly prepositions, however locationals such as perto 'near', diante 'in front', junto 'with', traz 'behind', riba 'on top', sometimes occur as postpositions, usually in combination with su(a) 'of' or with the possessive of pronouns. Thus, the following examples may be found.

North IP: mezinh dentr 'in the medicine', cuj pert 'near which'.

Daman: minh junt 'with me' (p. 519).

Mangalore: bureau's riba 'on the desk', nosso junto 'with us', su diante 'in front of [you?], bossa pert 'near your', minha traz 'behind me' (p. 902).

Cochin: sua perto 'near you' (p. 813).

Schuchardt also notes that this structure can be found in Cape Verde creole:

si diante 'in front of you(?)' (1883b p. 902).

But in all dialects na 'on, in', com 'with', de 'of', por 'by, for', vide 'because of', etc. are always prepositions.

A similar situation obtains in 19th century SLCP: sua perto 'near him', minha juntado 'with me' (p. 59), and even non-spatial onte anoite impé 'since last night' (p. 111), aquél per despois 'after that' (p. 60). But also perto ele 'near him' (p. 113), despois de casamento 'after marriage' (p. 108), tras de almári 'behind the wardrobe' (p. 111). These are probably not simply influences from SP, since in archaizing BP prepositions may also be found: ri:wə də ælmæ:rə = ælmæ:rə ri:wə 'on the wardrobe'. In 19th century SLCP the non spatials na 'on, in', per 'to, for, by', de 'of, for', vídè 'because of ', etc., are always prepositions, but in colloquial BP they are always postpositions and some have even become case suffixes.[9]

Although there were probably a few spatial postpositions in early IP, the generalization of postpositions in BP is clearly due to the influence of the local languages.

2.1.3 ADJECTIVES. In BT and BP adjectives always precede the noun they modify; in SP they may either precede or follow. In most of the other IP dialects the tendency is towards the order Adjective-Noun, but Noun-Adjective may be found in all dialects (except perhaps Diu). In 19th century SLCP too, Noun-Adjective is not entirely ruled out. But in BP this order occurs only frozen in songs and in compounds such as pa:nu#wæ:y 'rag' or di:yə#(g)ra:ndi 'Xmas-New Year's festival'. In excluding this order from

[9]See the discussion of case affixes in Section 2.2.

productive syntactic constructions of adjective and noun, BP conforms to the pattern of BT and Sinhala.

2.1.4 AUXILIARY. In SP verbal auxiliaries usually precede the verb, in BT they always follow. BP is somewhere in the middle, with some auxiliaries following the verb, and some preceding it. Among the former are tiñə 'had' in (5), te:m 'has', bota: 'put', (one of several completive aspect markers), and da: 'give', (an aspect marker denoting action performed for someone else's benefit); in the latter class are po:y 'can' in (6) and kerə 'want, intend'.

(5) a) BP ung gɔ:tə je:ntis-pə me: (ja:-) fəla:
[PPL]
a few people DAT EMPH PAS tell

PPL
BT oŕu koncam a:kkaL -ukku ta:n coll -i

BP tiñə
[PAS]
was

PAS CNC
BT iŕu-nt -a:ŕ

'He had only told a few people'.

cf. b) BP tinha fala-do a pouca gente somente
[3sg]
had tell PAS PPL to few people only

(6) a) BP əkə -ntu fu:lə na: po:y bota:
NEGPOT can

that LOC flower put

INF can NEGPOT

BT at -ila pu: po:T -a e:l-a:t

BP na:

TAG

BT -e:

'You can't put [i.e. sew] flower [designs] on that [sewing machine], can you.'

cf. b) SP naquilo não se pode botar fulas,
[INF]

on that NEG one can put flowers,

não é.

TAG [lit. NEG-is]

It is interesting that although BP po:y always precedes the main verb in affirmative sentences, in negative sentences a variant with infinitive followed by po:y has developed on the model of Tamil. Thus (7) would be a possible equivalent of (6) a); compare the actually recorded sentence in (8).

(7) BP(S)[10] fu:lə pə- bota: na: po:y

flower INF put NEGPOT can

[10](S) indicates that I have not verified this sentence with a consultant.

(8) BP tɔ:nə ækə midisa:m ində ungə-pə pə-
afterwards that measurement another one DAT INF

æra: na: po:y
miss NEGPOT can

'Henceforth, no one else can miss that measurement [because the surveyors have put in a marker].'

In all other varieties of IP as well as in 19th century SLCP the order is always Auxiliary-Verb as in SP.

2.1.5 RELATIVIZATION. Relative clauses (e.g. (4) c) k)) are imbedded sentences which can be derived from full sentences with equi-NP deletion of some kind. We take this fact to be a defining feature rather than a universal property. Relative clauses in BP and BT share several language-specific characteristics. In both languages structures giving rise to relative clauses are left-branching. Neither language has relative pronouns. There is no change in the word order of the embedded S. Any noun may be relativized, but all its associated case affixes and or postpositions are also deleted along with it. Finally there remains no surface trace of its role in the embedded S. For example, the deleted noun in the sentence underlying the relative clause in (4) c) would be in the locative case in both languages. In contrast, relative clauses in SP are right-branching; they have relative pronouns which are fronted; and traces of the role of the relativised NP in

the embedded S remain. The relative clause in (4) c) would be given as (9) a) or b) in SP.

(9) a) SP ao lugar em que (in which) eles se reúnem
to the place they REFL meet

b) SP ao lugar onde (where) eles se reúnem

In other varieties of IP and 19th century SLCP virtually all relative clauses follow the SP pattern. I have, however, found two instances of what I believe is an intermediate type:

10) a) North IP ós qui (REL) ja deu [PAS] anel ... (p. 160)
you PAS give ring

b) BP(S) bos ja: da: ænæ:lə ...

'The ring you gave.'

(11) a) 19thC SLCP Este dous pessão-s (PL) aquel dia que (REL)
These two people that day

b) BP (S) isti do:s pesa:m əkə di:yə ja:- (PAS)

19thC SLCP fazê comédè vos mistediá olhá
do folly you OPT-PAS see
fəyə kumæ:di bo:s mesti:(ñ)ə oya:[11]

'You should have seen the folly that these two people committed that day.'

---

[11] or mes(tə) oya: tiñə
OPT see PAS

The word order in the two examples follows the BP pattern; the only difference is that North IP and 19thC SLCP retain the relative pronoun.

2.1.6 LINKING PAST PARTICIPLE. The linking past participle (e.g. (4) c)) is a feature found in most of the languages of South Asia and is variously known as 'gerund,' 'absolutive,' 'indeclinable participle,' 'conjunctive participle,' and 'adverbial participle.' It is typically used to conjoin verbs in a temporal sequence which share the same subject and often other sentential elements as well. All but the last of the conjoined verbs appear as past participles. Two further examples of this construction are given below.

(12) BP əkələ:rə fəla: oya:
then tell see
[PAS] PPL
BT appa coll -i pa:ŕu

'Then having told [them] see [what ensues].' (I.e. Tell them and see).

(13) a) BP əwərə əla: | ində um luwa:rə tudu
now there another a place all
LOC
BT ippa anka | inn -oŕu iTatt -ila ella:m

BP rəspa:
scrape
[PAS] PPL
BP curaNT-i

'Now there | in another place having [had] all scraped.'

b) BP ki: se: ti:ntə onta: -tu

what COND paint paint PFC

BT enna eNT -a:l -um may pu:c -i [-ŕŕu][12]
QUOT even PPL

BP tə -tira:
PRES

take

eTu-kk-r -an
PRES CNC

'Having painted on whatever paint, I am taking it.'

(I.e. 'In another place I am having it all scraped down and painted with some kind of paint or another and getting it back.')

These formations sound as unnatural in SP as they do in English. There is no distinct morphological marking for the past participle in BP, but in both BT and BP the participle may be marked with the perfective suffix BP -tu/ BT -ŕŕu (as in (4) c) and (9) b)) with little change in meaning. The perfective past participle is a distinct form in both languages.

The linking past participle does not appear to be found in any of the other IP creoles, nor in 19th century SLCP, though the following example from Mangalore is a good candidate.

---

[12]Omitted from the original BT translation.

(14) Mang. E levanta ele, ja vi papa 's pert
and get up he PAS come father GEN near
'And he got up and came to his father.'

2.1.7 QUOTATIVE. The quotative construction, which is of Dravidian origin and has no parallel in SP, is used to mark material which is the object of an overt or implied verb of thinking, naming, speaking, etc. The quoted material is followed by the quotative particle, BP fala:, BT en, both formally verbs meaning 'say, tell,' though in BT en has become specialized in its quotative function and is no longer used as a main verb. Quotative constructions are not formally different from other embedding phenomena: the quotative particle most commonly occurs as a linking past participle (e.g. (2) d), (4) h)) and it is sometimes also found as a verbal adjective in a relative clause formation.

The following excerpt from a taped conversation is particularly rich in quotatives. Note that the verb fala: functions sometimes as the quotative particle and sometimes as the main verb 'tell, say.' There is unfortunately no BT translation of this conversation.

(15) BP e:w e:li-ntu ja:-fəla: əwərə ung suma:nə
I him LOC PAS tell now one week [contrastive intonation]
A 'It is now a week since I told him [to tell you to come and see me].'

(16) BP oy me: ja:-fəla:

today EMPH PAS tell

B 'He told [me] just today.'

(17) a) BP tɔ:nə e:w ɔ:ntə-otər-di:yə

afterwards I the day before yesterday

ja:-punta:

PAS ask

A 'Then the day before yesterday I asked [him]:'

b) BP bo:s ja:-fəla: wo: ricəd -ntu a: fəla:-tu

you PAS tell INDEF Richard LOC at QUOT PFC

A 'Did you tell [him]?-- Richard.'

c) BP e:li ja:-fəla: e:w ja:-ka:-fəla: fəla:-tu

he PAS tell I PAS PFC tell QUOT PFC

A 'He said he [had] told [you].'

d) BP inna: bo:s tə -fəla: əwərə

here you PRES tell now

oy me: ja:-fəla: fəla:-tu

today EMPH PAS tell QUOT PFC

A 'Now here you are saying he told [you] just today!'

(18) BP ja:-fəla: wi:də me: ja:-wi: istɔ:ndə
PAS tell because EMPH PAS come this way

B 'It is precisely because he told me that [I] came over here.'

2.1.8 CONDITIONAL. In BT the conditional is formed by adding the suffix -a:(l) to the past tense stem of the verb. In BP too, the usual conditional is formed by postposing the conditional element se:/sərə 'if' to the verb, which may be optionally past or tenseless[13] (e.g. (3), (4) a), c), d)). The conditional of SP, other IP creoles, and 19th century SLCP is signalled by the conjunction se 'if' (also written si in some creoles), which always precedes the verb, and usually comes at the beginning of the clause.

2.2 Morphology

As might be expected given its pidgin origins, BP has a much less complex morphology than BT. For example, BT verbs are marked for subject concord but BP verbs are not; tense/aspect/mode markers are more fully integrated into the verb in BT than in BP, where they may often be treated as particles or clitics; and finally, the BT forms exhibit many more morphophonemic alternations.

Disregarding the fact that a given semantic or syntactic

[13] se:/sərə may also follow other syntactic categories and is therefore not treated as a suffix. BP also has a less frequently used morphological conditional formed by prefixing kəm(də) (<SP quando) to the verb stem. Blended forms with both the prefix kəm(də) and the post-clitic se:/sərə are heard occasionally.

category may be represented by a morphological marking in one language and a postposition or particle, or even a combination of morphological and other markings, one finds quite a close correspondence between the sets of categories which must be set up for the two languages. Moreover, the details of the surface representation too are often surprisingly close. We shall illustrate the situation by outlining the inflection of nouns. Verb morphology, which is more complicated, and not so well understood, will only be touched on.

2.2.1 NOUNS. Nouns in BT and BP inflect for number and case. In SP they inflect only for number.

i) Number

Singular is morphologically and semantically unmarked in BT, SP and in all the IP creoles. The plural suffixes are BT -ka(L)/-a:ŕ; SP, IP -s (e.g. (2) c), (4) f)). In BT and BP, as in other varieties of IP, a noun preceded by a quantifier is usually not overtly marked for plurality, (e.g. (4) d)).

ii) Case

I shall first give the BP and BT data, and then discuss the development of each affix separately.

Case inflection in SP and BP is by means of suffixes, which follow the plural marker if any. The nominative is morphologically and semantically unmarked. In BT many nouns have a special stem to which the (non-nominative) case

suffixes are added in the absence of the overt plural marker. (19) illustrates the oblique stem in BT and the order of affixes in both languages. Note that the -ka/-kaL alternation in the plural suffix in BT is phonologically conditioned: the former variant occurs before a consonant or pause, the latter before a vowel.

(19) a) BT ma:Tu
cow [NOM]
BP ba:kə

b) BT ma:Tu -ka
cow [NOM]-PL
BP ba:kə -s

c) BT ma:TT-ukku
cow -DAT
BP ba:kə-pə

d) BT ma:Tu-kaL-ukku
cow -PL -DAT
BP bə:kə-s -pə

In many South Asian languages it is difficult to distinguish objectively between case suffixes and postpositions. In BT case suffixes are added to an oblique stem, while postpositions are not. In BP we have used the criterion that case suffixes are stress neutral and postpositions are not.

The case affixes for BP and BT are given in Table 1. In BP the accusative suffix is -pə for humans, -Ø for inanimates, and either -Ø or -pə for non-human animates. In BT -a(y) usually occurs with [+definite] nouns and -Ø with [-definite] nouns. [+human] nouns are nearly always [+definite], and with [-human] nouns the suffix -a(y) may often be omitted. The BP genitive -su has an unreduced

Table 1. BP and BT Case Inflection

	BP	BT
nominative e.g. (2)a)b)d), (3),(4)d)d)f)j)j)*	ø	ø
accusative e.g. (1), (4)f)g)h)h)	ø / -pə	ø / -a(y)
dative e.g. (1), (2)b)	-pə	-ukku
genitive e.g. (1),(4)j)	-su(wə)	-ŕa
locative e.g. (2)a), (3), (4)i)k)	-ntu	-ila(y)/-iTTa(y)
associative	see text	-o:Ta(y)
instrumental	see text	-a:la(y)

*Examples listed twice contain two instances of the form in question

variant -suwə which is rarely used. The BT locative suffix -ila(y) is used with inanimates; -iTTa(y) occurs with animates. In BP, when no postposition occurs following the locative of a human noun,-a: may be optionally added. Although -nt-a: is historically from SP juntado 'joined', this a: is best treated as a postposition in BP, since it may be stressed and does not co-occur with other postpositions. The BP equivalent of BT associative and instrumental cases are the postpositions ju:ntu and wɔ:ndə respectively.

Merely listing corresponding forms as we have done does not, of course, prove that the systems are congruent. A thorough proof of this would require an examination of the entire range of usage of each of the forms to find out to what extent their distributions match. One does find points of non-parallelism; the differential usage of the overt accusative just mentioned is one example, another is the use of accusative in BP and dative in BT to signal the goal of a verb of motion as in (3), and (4) k). In the main however, one may set up correspondences between BT and BP forms which hold over a wide variety of uses.

The occurrence of case suffixes is perhaps the most striking feature of BP morphology. Historically these have all developed from postpositions which have become reduced and stressless. In consciously archaizing BP prepositional constructions also occur. In place of N_1 -su N_2 we can find N_2 də N_1 as in SP. This construction still survives in colloquial BP in frozen forms such as bɔ:wər#d#a:w 'water-

melon' or pa:nu#də#mæ:zə 'table cloth' (= mæ:zə-su pa:nu). The periphrastic genitive, however, is a marginal construction in BP, as can be seen by the fact that the order may be reversed with da preceding the head noun N_2 as in (20) b).

(20) a) aro:s də məntε:yə (rice of butter) 'ghee rice'

b) məntε:yə də ərɔ:s id.

The dative and locative (as well as postpositional phrases) may also be replaced by prepositional phrases with either nə or də and more rarely pə(r) and kum. The seemingly indiscriminate substitution of prepositional phrases for case suffixes is illustrated by (21) - (23), all sentences elicited from the same speaker in response to English cues. The reader will note other non-colloquial features, such as SVO word order in these examples.

(21) a) BP e:w mes(tə) pəpiya: kum prensipa:l

I must speak with boss

'I must speak with the boss.'

b) BP e:w mes(tə) pəpiya: pər prensipa:l id.

c) BP e:w mes(tə) pəpiya: nə prensipa:l id.

d) BP e:w mes(tə) pəpiya: prensipa:l id.

cf. e) BP e:w mes(tə) pəpiya: prensipa:l-ntu id.
LOC

(22) a) BP bəsi:m ja:-kəy nə ca:m
plate PAS fall on floor
'The plate fell on the floor.'

cf. b) BP bəsi:m ja:-kəy ca:m-ntu id.
LOC

(23) a) BP miñə muye:rə ja:-limpa: nə ispe:y pa:nu wɔ:ndə
my wife PAS clean mirror cloth with
'My wife cleaned the mirror with a cloth.'

cf. b) miñə muye:rə ja:-limpa: ispe:y pa:nu wɔ:ndə id.

Let us now look at each of the BP case affixes individually and find their correlates in other IP dialects and in 19th century SLCP.

The BP ACC/DAT -pə has developed from the SP preposition pera (modern para) 'to, for'. Other dialects of IP and 19th century SLCP retain a prepositional construction with par, pa, per, pê, pu, pum, pur, por, pr or pro, and as in BP this is usually used to mark indirect objects and [+human] direct objects. As all SP dialects agree in this usage it probably arose early in IP, very likely under the influence of the northern Indo-Aryan languages. Another factor in its history may have been the so-called 'prepositional accusative' of SP, whereby human objects could at one time be marked with the preposition a 'to'. This usage has since died out in SP, except when the object is Deus 'God'. The SP preposition a overlaps semantically and

distributionally with para; in particular, both can be used to mark the indirect object as in (24).

(24) SP Eu mandei uma carta $\left\{\begin{matrix}\text{a} \\ \text{para}\end{matrix}\right\}$meu pai.

I sent a letter to my father.

The variant with para, though less common in Portugal, is the norm in Brazil. In IP, virtually all the functions of a have been taken over by para and its variants, and in some dialects (BP, for one) a does not survive at all. Presumably then, the a of the prepositional accusative also yielded to para and became reinforced as a marking for human indirect objects by the influence of indigenous languages. The fact that the accusative of human nouns is specially marked in BP can thus not be seen as a result of BT influence. However, since in other forms of IP, including 19th century SLCP, para and its variants are always prepositions, the fact that in BP it has been postposed and reduced to a case suffix can be ascribed to BT influence.

The postposed genitive su, sua, etc. is found in all the other IP dialects, though in all but Mangalore the SP construction with de is more common. In Mangalore (and Bombay) su has been reduced to -s and is written 's as in English. In 19th century SLCP N_1 -su N_2 is more common than N_2 de N_1 (Dalgado p. 58). We shall simply assume that postposed su(a), etc. may be treated as a case suffix in all dialects, though the fact that in all the above

dialects su(a) also exists as a third person possessive pronoun perhaps favors a syntactic analysis in all but Mangalore, where reduced -s is found. Whatever the correct analysis may be, it is clear that the postposed su(a) genitive must have been found in early IP[14] in competition with the SP expression with de. The fact that the prepositional genitive has been dropped in favor of the -su genitive in BP may be attributed to indigenous influence.

The BP locative suffix -ntu is ultimately from SP junto 'joined'. In most of the IP creoles expressions such as $\left[\begin{smallmatrix}N\\+human\end{smallmatrix}\right]$ su junto or $\left[\begin{smallmatrix}N\\-human\end{smallmatrix}\right]$ junto 'with N' can be found. These also exist in BP, but a reduced form of the postposition -ntu has become specialized as a locative case affix with the meaning 'at, in, on, etc.' Examples of the postposition and case suffix in BP are given below.

(25) əkə ju:ntu əki: tə- sirwi: ta:m
that with here PRES work also
'In addition to that, they are working here also.'

(26) tɔ:nə fuwa:m-ntu gərda: mestə-kuzə
then fire LOC keep OPT boil
'Then you must put [it] on the fire and boil [it].'

---

[14] It also occurs in other Asian Portuguese creoles (Singapore, Malacca, Java and Macau) as well as in Papiamentu. (Cf. Hancock (1975) pp. 222-3.

(27) nɔ:ywə ta:m əkə ley me:
bride also that like EMPH

nɔ:ywə-su je:ntis ju:ntu lo- ənda: gre:jə
bride GEN people with POT go church

'The bride also in the very same way goes with the bride's people to the church.'

(28) e:w ja:-fəla: ənda: pa:y -ntu fəla:-tu
I PAS tell go father LOC tell PFC

vira: vi: miñə-ntu diñe:ru tri:yə fəla:-tu
back come I[15] LOC money bring QUOT PFC

'I said, "Go tell your father [then] come back and bring me the money."'

2.2.2 VERBS. In noun morphology the two languages exhibit a neat pattern of formal and semantic congruence. In verb morphology, on the other hand, we find that the languages are formally not very similar. While BP, makes use of both prefixes and suffixes, BT has only the latter. BT makes more formal distinctions than BP; subject concord, found in BT but not in BP, is a case in point. Moreover, the individual morphemes making up the larger forms do not always match up as well as they did with nouns. Nevertheless, on the semantic level, we again find that systematic corres-

[15]The pronominal declensions are irregular; here the LOC affix is added to the GEN stem, cf. bosə-ntu 'to you' nosə-ntu 'to us' but eli-ntu 'to him' botus-ntu 'to you [HON]'.

spondences can be set up between the morphologically complex forms, which hold through a wide range of usage. Tables 2a and 2b constitute a partial listing of such correspondences. We have chosen two regular verbs, BP prendə 'study, learn' and BT paTi id., and have assumed a first person singular subject where possible.

Comments:

By independent forms are meant those which may occur as the main verb of a matrix sentence; all others are dependent forms.

For all verbs in BP and most verbs in BT the verbal base is used as an imperative.

In BT the present and potential (often called 'future') have different formations for [-human] subjects, which are omitted here for simplicity.

BT -kk-, which appears in the present and elsewhere, may be treated as an augment which occurs with certain transitive verbs; cf. paN 'do, make' in (4) f), a transitive verb which does not form the present stem with -kk-.

In formal Tamil, perfective is a periphrastic construction composed of the past participle and an inflected auxiliary; this has become morphologized in most colloquial dialects.

We have mentioned above that the BP conditional form which more closely resembles that of BT, i.e. the form with the postposed conditional element se:/sərə, is by far the more common.

Table 2a. Some BP and BT Verb Correspondences: Independent Forms

Category	BP	BT	Gloss
imperative	prendə [BASE]	paTi [BASE]	'study!'
present (e.g. (2)b),(4)f)g)h))	tə- prendə PRES	paTi-kk-r -an [+HUM] PRES CNC	'I am studying'
past	ja:-prendə PAS	paTi-cc -an PAS CNC	'I studied'
potential	lo- prendə POT	paTi-pp -an POT CNC	'I would study, will study, study, can study, etc.'
perfective PAS	ja:-ka:-prendə PAS PFC	paTi-cc -iˊ -fˊ -an PAS PFC PAS CNC	'I have studied, studied [and finished], etc.'
perfective POT	lo- ka:-prendə POT PFC	paTi-cc -iˊu-v -an PAS PFC POT CNC	'I will (would) study [and finish]'
descriptive	ki- (tə)-prendə DESC PRES	paTi-kk-r -a -(na:n) PRES DESC CNC	'I study, used to study'
permissive	jesə-prendə [Q INT] PER	paTi-kk-a -TT(um) a: INF PER Q	'should I study?' 'may I study?'
negative (e.g. (1))	nuku-prendə NEG	paTi-kk-a -illa(y) INF NEG	'I am not studying, didn't study'
volitive-negative (e.g. (4)i))	na: prendə NEGPOT	paTi-kk-a ma:TT -an INF VOLNEG CNC	'I won't study, don't study, wouldn't study'
negative-imperative	numis(tə)-prendə NEGIMP	paTi-kk-a:ta NEGIMP	'don't study'

Table 2b. Some BP and BT Verb Correspondences: Dependent Forms

Category	BP	BT	Gloss
conditional (e.g. (3), (4)a)d)e))	(ja:)-prendə{se: / sərə PAS COND kəm- (ja:)-prendə COND PAS	paTi-cc -a:(l) PAS COND	'if I study, studied'
infinitive	pə(r)-prendə INF	paTi-kk-a INF	'to study'
past participle	prendə [BASE]	paTi-cc -u PAS PPL	'having studied'
perfect participle (e.g. (4)c))	(ja:)-prendə-tu PAS PFC	paTi-cc -iŕ -ŕ -u PAS PFC PAS PPL	'having studied [and finished]'
verbal adjective (PRES ((4)c)k))	tə- prendə PRES	paTi-kk-r -a PRES DESC	'studying; who studies; etc.'
verbal adjective (PAS)	ja:-prendə PAS	paTi-cc -a PAS DESC	'studied; who studied; etc.'
verbal noun	ki- (tə)-prendə DESC PRES	paTi-kk-r -a -(tu) PRES DESC N	'studying'

The permissive is restricted to third person subjects in affirmative sentences and to first person subjects in interrogative sentences. It does not occur in negative sentences.

The infinitive also occurs as an independent form with first person subjects in interrogative sentences with the same meaning as the permissive.

The present verbal adjective, verbal noun and descriptive aspect are closely related in both languages. The latter two are identical in BP. In BT pronominal elements are suffixed to the verbal adjective to form the verbal noun and descriptive; these may even be deleted under certain conditions. The situation is in keeping with fact that BT has a richer morphology than BP and makes formal distinctions which BP does not. Note that despite any formal similarity the three forms are readily distinguishable syntactically. The descriptive is an independent form, and the other two are dependent. Of these, the verbal adjective always occurs with an overt noun or pronoun which it modifies, and the verbal noun is always imbedded under an NP node.

The negatives of the two languages need further study, and hence only three negative forms have been included in the above list. Even from the scanty evidence at hand, it is is clear that they belong to their own semantic system and do not closely correlate with the non-negative forms.

One further observation needs to be made. In BP there is

no formal distinction between negative potential and volitive negative; in BT the volitive negative is restricted to [+volitive] surface structure subjects, and the negative potential to [-volitive] surface structure subjects. The BT negative potential occurs in (1), where unfortunately it is equivalent to the BP simple negative nuku because the verb teŕi is somewhat aberrant. Cf. (6) a).

The purpose of this discussion has been to show that despite surface differences in the morphological make-up of the verb systems of BT and BP, there is substantial congruence in the structure of the semantic distinctions which each system exploits. I shall not attempt to investigate the verbal systems of other IP dialects, though I think that if sufficient data were available such an investigation would show that while there is substantial formal similarity to the verbal system of BP, the categories of the other IP systems do not match those of the BP system.

2.3 Conclusions

We have dealt with a number of syntactic and morphological phenomena which indicate that the indigenous languages of Sri Lanka have had a great deal of influence on Batticaloa Portuguese. In each case BP is closer to the BT model than SP, other IP dialects, or even nineteenth century SLCP. Moreover, evidence from archaizing BP seems to indicate that the influence has been quite recent. A summary of the results of the above investigation appears in Table 3.

Table 3. Some Grammatical Phenomena Showing Indigenous Influence in BP

	BT	BP	SLCP 19thC.	Cochin	Mang.	North	Daman	Diu	SP
Syntax									
SOV in non-copulative sentences	+	+	±	-	(-)	(-)	(-)	(-)	-
Postpositions	+	+	(-)	(-)	±	(-)	(-)	-	-
Adjective precedes noun	+	+	(+)	±	±	(+)	±	+	±
Auxiliary follows verb	+	±	-	-	-	-	-	-	-
Left-branching relative clause	+	+	(-)	-	-	(-)	-	-	-
Linking past participle	+	+	-	-	-	-	-	-	-
Quotative	+	+	-	-	-	-	-	-	-
Conditional element follows verb	+	(+)	-	-	-	-	-	-	-
Noun Morphology									
ACC/DAT suffix	+	+	-	-	-	-	-	-	-
GEN suffix	+	+	(+)	(-)	(+)	(+)	?	(-)	-
LOC suffix	+	+	-	-	-	-	-	-	-

Key: + always
(+) usually
± no apparent trend
(-) rarely
- never
? no data available

With respect to any particular phenomenon it is usually possible to argue that the situation in BP represents a natural development from that in IP, and hence that it is not necessary to invoke outside influence since such a development could have taken place independently. However, we are not in the presence of a single phenomenon, but a cluster of phenomena, in all of which development in BP has been towards BT. The case for external influence is therefore a strong one.

3 Phonology

In this section we shall deal with the phonological aspects of indigenous language influence in Batticaloa Portuguese. Since the early works on Indo-Portuguese are based largely on written materials, they are not as useful here as they were in morphology and syntax. Even so, Dalgado spent a year in Sri Lanka and his work on Sri Lanka Portuguese (1900) contains some useful descriptions of actual pronunciation. To this may be added the few comments of Vasconcellos (1901 pp. 144-148) based on his chance encounter with two mother-tongue speakers of SLCP from Colombo at the Paris world's fair in 1900.

3.1 Phonological inventory

The phonological inventory of BP was given in Tables 1 and 2 of Chapter II. Here for comparison we give the (autonomous phonemic) inventories of 16th century standard

Portuguese (Tables 4 and 5)[16] and of Batticaloa Tamil (Tables 6 and 7). As I shall be presenting some phonological rules of Tamil, a distinctive feature analysis of the Tamil inventory is given in Table 8. No major rules of SP will be presented here.

We shall now examine in detail some of the areas in which BT and SP differ in order to evaluate the phonological influence the former has had on BP. We shall look at contrastive vowel length, vowel nasalization, some vowel allophones, the stop series, the palatals and the retroflexes.

3.2 Vowel length

Both BP and BT have contrastive vowel length and predictable stress. In SP the reverse situation obtains: stress is not fully predictable and vowel length is not significant. BP is thus typologically similar to BT rather than SP. None of the studies on other dialects of Indo-Portuguese state explicitly that vowel length is contras-

---

[16] Cf. Hart, Carvalho (1962). The vowel inventories of 16th century and modern SP are the same. The only development in the consonant inventory has been the change of the affricates to their corresponding spirants. I.e. /t͡s s/ > /s/, /č š/ > /š/, /d͡z z/ > /z/, and /ǰ/ > /ž/. The old contrasts have been retained in the writing system, however: /t͡s/ = ç before back vowels and c before front vowels, /s/ = ss intervocalically and s finally, postconsonantally, and before voiceless consonants, /d͡z/ = z, and /z/ = s intervocalically and before voiced consonants. Hart assumes j was already pronounced [ž] by the 16th century, but I have taken it to be still an affricate, since it shows up as such in all the Portuguese creoles (except Principense). I have not posited glides for SP as [y w] may be treated as allophones of /i u/.

Table 4. Consonants of 16th C. Portuguese

		bi-labials	labio-dentals	dento-alveolars	palatals	velars
stops	vl.	p		t		k
	vd.	b		d		g
affricates	vl.			t͡s	č	
	vd.			d͡z	ǰ	
spirants	vl.		f	s	š	
	vd.		v	z		
nasals		m		n	ñ	
med. res.				r r̃*		
laterals.				l	l̃	

*r̃ may be treated as /rr/

Table 5. Vowels of 16th C. Portuguese

	front	back
high	i ĩ	u ũ
mid	e ẽ	o õ
low	ɛ	a ã ɔ

Table 6. Consonants of BT

	bi-labials	labio-dental	dental	alveolars	post-alveolar	retro-flexes	pala-tals	velar
stops	p		t			T*		k
affricate							c	
nasals	m			n		N*	ñ	
median resonants				ṛ	r			
lateral liquids				l		L*		
glides		v**					y	

* I have used capital letters in phonetic transcriptions of BP to indicate post-alveolar, but non-retroflex consonants. In phonemic transcriptions of BT I am following convention and using capitals to represent retroflex consonants. In phonetic notation a retroflex will be symbolized by a dot under the appropriate lower-case letter. So, e.g. BT /N/ is realized as [n̩].

**/v/ is the usual representation in romanization of Tamil வ். It is a frictionless continuant, with the same variants as BP /w/.

Table 7. Vowels of BT

	front		central		back	
	short	long	short	long	short	long
high	i	i:			u	u:
mid	e	e:	a*		o	o:
low				a:		

*/a/ is the conventional romanization of Tamil அ, the short counterpart of ஆ (/a:/). Phonetically it is [ə].

Table 8. Distinctive Feature Composition of Batticaloa Tamil Phonemes

	i:	i	e:	e	a:	a	o:	o	u:	u	y	v	ŕ	r	l	L	p	m	t	n	c	ñ	T	N	k
Syllabic	+	+	+	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Sonorant	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	+	–	+	–	+	–	+	–
Consonantal	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+	+	+
High	+	+	–	–	–	–	–	–	+	+	+	–	+	–	+	–	–	–	–	+	+	+	–	–	+
Back	–	–	–	–	+	+	+	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+
Low	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Anterior	–	–	–	–	–	–	–	–	–	–	–	–	+	–	+	–	+	+	+	+	–	–	–	–	–
Coronal	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	–	–	+	+	+	+	+	+	–
Round	–	–	–	–	–	–	+	+	+	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Long	+	–	+	–	+	–	+	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Nasal	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	+	–	+	–	+	–
Lateral	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–
Continuant	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	–	–	–	–	–	–	–	–

tive. As the difference between long and short vowels would not appear in the transcriptions of the time and since the writers in most cases had never heard the languages spoken, they could hardly have had the means to discern contrastive length. Nevertheless, in Dalgado's study of Daman Portuguese there is the enigmatic statement, 'There are no mute vowels like a, e, o in Indo-Portuguese: they are long or short.'[17] The reader will need to know that in all dialects of Portuguese, including the creoles, earlier unstressed /a e o/ have reduced finally and, depending on the dialect, in some or all medial environments. These reduced vowels, in continental Portuguese especially, have a tendency to drop out altogether. Dalgado is probably saying, then, that in Indo-Portuguese these reduced vowels are not deleted. It is the second part of the sentence, 'they are long or short', which is difficult to interpret. I do not think he can be saying that there is contrastive length, since he gives no examples and never indicates vowel length in his transcriptions. He is probably referring to the fact that in many dialects of Portuguese stressed vowels are longer than unstressed vowels. Formally, these dialects have the low level rule of vowel lengthening, given in Chapter II (rule (318)). Dalgado's statement should probably be interpreted to mean that the vowels written as a, e, o are never deleted as in SP, but are either long (when

---

[17] 'Não ha vogaes surdas no indo-português, como a, e, o: são longas ou breves.'

stressed) or short (when unstressed). But the possibility always remains that he is, in fact, reporting that Indo-Portuguese has contrastive vowel length. There is, however, strong evidence that in early SLCP the original SP rule of vowel lengthening was still in force, and hence that vowel length was not functioning contrastively. The evidence comes from a class of loans from Sinhala (S) and Tamil (T) in which original stressed short vowels have been lengthened in the creole. Some examples are given below.[18]

(29)	BP [bá:ṇdu]	'sp. of insect'	cf. BT	[véṇdu]	id.
(30)	BP [pʒ́:ṭi]	'box'	cf. BT	[pʒ́ṭṭi]	id.
(31)	BP [tá:yru]	'curd'	cf. BT	[téyiṛ]	id.
(32)	BP [né:li]	'paddy'	cf. BT	[néllu]	id.
(33)	BP [sú:rə]	'toddy'	cf. S	[súrə]	id.
(34)	BP [ká:ndiyə]	'Kandy'	cf. S	[kándə][19] 'mountain'	
(35)	BP [gó:rkə]	'goraka'[20]	cf. S	[górəka]	id.

We may assume then that early SLCP had significant stress

---

[18] It is possible that (29) - (32) were actually borrowed from Malayalam, a Dravidian language closely related to Tamil and spoken on the southwest coast of India, where the Portuguese possessions of Cannanore, Mahé, and Cochin were located. Cf. Malayalam vaṇṭu (DED 4285), peṭṭi (DED 3600), tayir (DED 2517), and nel (DED 3112). If this is the case, these words may have already been present in the Indo-Portuguese brought to Sri Lanka. This cannot of course be true of the loans from Sinhala.

[19] The Sinhala name for the Kingdom of Kandy was [kándə úḍə ráṭə] 'the country on the mountain'. The Creole word may show influence from Sinhala [káṇḍiyə] 'hillock'.

[20] Garcinia cambogia guttiferæ, a sour fruit used in curries.

and predictable vowel length. Eventually length came to be interpreted as the significant feature and stress as predictable. Once this reinterpretation had been made, short stressed vowels were free to develop, which would contrast with the more common long stressed vowels. These short stressed vowels came from inherited words which usually occurred in unstressed environments, from loans from Dutch and indigenous languages, and from phonological developments within the creole.

Under normal conditions of development a Standard Portuguese word would appear in SLCP with a long vowel in place of the SP stressed vowel.[21] However, there are examples such as (36) - (44) of SP words which appear in SLCP with a short stressed vowel instead of the expected long vowel.

(36) BP /eli/ 'he' cf. SP ele id.

(37) BP /ælə/ 'she' cf. SP ela id.

(38) BP /botus/ 'you PL/HON' < SP vós outros 'you others'

(39) BP /etus/ 'they HON' < SP elis outros 'they others'

(40) BP /isti/ 'this' cf. SP este id.

(41) BP /miñə/ 'my, mine' cf. SP minha id.

(42) BP /nosə/ 'our, ours' cf. SP nossa id.

(43) BP /bosə/ 'your(s)' cf. SP vossa id.

(44) BP /ung/ 'one, a' cf. SP um id.

---

[21] Dalgado notes that in 19th century SLCP 'the tonic vowels are, as a rule, longer than in Portugal . . .' (1900 p. 26). Though clearly by this time vowel length must have been contrastive, Dalgado does not mention that this is the case.

These are all words which could be expected to be unstressed in most of their occurrences in the context of discourse.[22] As lengthening applied only to stressed vowels, these words would in most of their occurrences have short vowels. Consequently, either they were assigned short vowels on the basis of their most common phonetic shape at the time when stressed vowels were reinterpreted as long vowels, or they were initially assigned long vowels and then subsequently phonologically restructured. As might be expected, this was not a systematic development. Although some pronouns, such as (36) - (40), turn up in BP with short vowels, others such as (45) - (48) have long vowels.

(45) BP /e:w/	'I'	cf. SP eu		id.
(46) BP /bo:s/	'you'	cf. SP vós		id.
(47) BP /no:s/	'we'	cf. SP nós		id.
(48) BP /osi:r/	'he/she HON'	cf. SP o senhor		'he/you HON'

Dutch and the indigenous languages are the source of a number of words with short stressed vowels:

(49) BP /blət/	'tin'	cf. D blat		'leaf, sheet'
(50) BP /gɔrgəl/	'throat'	cf. D gorgel		'throat, neck'
(51) BP /stəm/	'stem'	cf. D stam		id.
(52) BP /krel/	'curl'	cf. D krul		id.

[22] The development of a short vowel in BP /oy/ 'today' < SP hoje id. is probably a sporadic innovation, motivated perhaps by the need to keep this item distinct from BP /o:y/ 'eye' < SP olho id.

(53) BP /slæktə/ 'bad' cf. D <u>slecht</u> id.

(54) BP /rɔl/ 'roll' cf. D <u>rollen</u> id.

(55) BP /ceku/ 'oil press' cf. T /cekku/ id.

(56) BP /cəlli/ 'money' cf. T /calli/ id.

(57) BP /vəhupu/ 'subject' cf. T /vakuppu/ id.

We may assume that at the time these items were borrowed the vowel lengthening rule had been lost.

Finally, short stressed vowels also arise from phonological changes within the creole. Sometime after stressed vowels were reinterpreted as long, in BP there was a shifting of stress from the final syllable of verbs in /-e:/ (< SP verbs in /-ér/) to the vowel of the immediately preceding syllable. This was accompanied by the reduction of final /e:/ to /ə/.[23] In all Portuguese creoles, the final -<u>r</u> of SP infinitives was lost. Thus SP -<u>er</u> > creole /-é/, usually written -<u>ê</u>, which in SLCP would have become /-e:/, though in the 19th century documents this is still written -<u>ê</u> or -<u>é</u>. Stress shifting in BP must have taken place after the lost of the vowel lengthening rule, since in bisyllabic stems there has been no lengthening of the initial syllable vowel to which stress was shifted. In the examples below some BP bisyllabic stress shifting verbs are compared with their counterparts in 19th century SLCP and SP. The earlier SLCP forms are taken, unless otherwise

[23]Since BT has no final syllable /e:/ except in monosyllables, it is possible that its influence had something to do with this development in BP, as it results in a parallel distributional constraint.

indicated, from Dalgado 1900 (Dal), Callaway 1823 (Cal), and Fox 1819.

(58) BP /kerə/ 'what' [kérə]

cf. Dal querré

Cal que´re[24]

Fox quere

SP querer

(59) BP /murdə/ 'bite' [múrdə]

cf. Dal murdé, murdá,[25] mourdê

Cal mur´de

Fox murdue

SP morder

(60) BP /fəyə/ 'do, make' [féyə]

cf. Dal faé

Cal fai, fa´ze

Fox fai, faze

(61) BP /disə/ 'get down' [dísə]

cf. Dal disce, descí

Cal dis´ce

SP descer

---

[24]Callaway's stress indications cannot, in most cases, be factual, because they are often at variance with both SP and BP.

[25]Verbs in BP nearly always reflect the conjugation class of their SP etyma; hence it is difficult to believe that the final vowel was ever unstable, as this and numerous other examples suggest.

With -e: verbs of more than two syllables, stress shifted to the penultimate syllable vowel and this was reinterpreted as a long vowel.

(62) BP /risi:bə/ 'receive' [risí•bə]

cf. Cal re´cebe

Fox recebe

SP receber

(63) BP /inti:ndə/ 'understand' [intí•ndə]

cf. Dal intindê

Cal intin´de

Fox entende

SP entender

(64) BP /uskundə/ 'hide' [uskú•ndə]

cf. Dal uscundê, escundê

Cal escon´də

Fox esconde

SP esconder

(65) BP /iskruwə/ 'write' [iskrú•wə]

cf. Dal eskruvê

Cal es´cruvi

Fox escrovi

SP escrever

To what can we attribute the lengthening observed in these examples? Certainly not to the operation of the original

vowel lengthening rule, unless stress shifting took place earlier in polysyllabic -e: verbs than in bisyllabic -e: verbs. A more appealing interpretation is that the lengthening came about because it was impossible to maintain a distinction between long and short non-initial syllable stressed vowels. As was explained in Chapter II, underlying long vowels in non-initial syllables which are assigned stress appear on the surface as half-long and in checked syllables may even be short. It would not be surprising, then, that these vowels should have fallen together with the new non-initial syllable stressed vowels introduced by stress shifting. Although stress shifting can be observed most commonly in -e: verbs, where it was exceptionless, it can also be found sporadically in other items.

(66) BP /cəpə/ 'hat' [čə́ppə ~ čə́pə]

cf. LV[26] chappé

Dal chapé

Cal chap'pe

Fox chape

[26][Leite de] Vasconcellos 1901.

(67) BP /əkə/ 'that' [ákkə ~ ákə]

cf. LV ácca

Dal aquél

Cal a'quel

Fox aquel

SP aquele

Stress shifting in BP appears to be a fairly recent phenomenon. Older speakers in archaizing speech occasionally produced stressed final [é·], as in (68) - (69).

(68) BP /mure: ~ murə/ [muré· ~ múrə] 'die'

(69) BP /kume: ~ kumə/ [kumé· ~ kúmmə] 'eat'

Final [é·] also survives in songs:

(70) səwa:m kerə puyər kerə

soap need powder need

ba:rwə pə- truse:

moustache INF twist

ænæ:lə kerə dɔ:ti kerə

ring need dowry need

buni:tu per pərse:

beautiful to appear

'You need soap and you need powder
to twist your moustache
You need a ring and you need a dowry
To look beautiful.'

(Sung as part of the wedding celebrations when the bride and groom are getting up.)

We have seen that the development of contrastive vowel length took place after the establishment of Portuguese in Sri Lanka and that the functional load of this opposition has been gradually increasing.

Sinhala, Tamil and Dutch all have contrastive vowel length and while the development of this feature in SLCP is clearly a result of external influence it is not possible to decide whether Dutch is to be included among the influencing languages. If when the Dutch arrived on the scene Vowel Lengthening was still in operation we might expect that some early Dutch loans into BP would exhibit the lengthening of initial syllable short vowels as with the Sinhala and Tamil loans in (29) - (35). That no examples can be found is perhaps an indication that vowel length was already contrastive when the Dutch arrived.

Finally it should be noted that although BP, BT and S are typologically similar in having contrastive vowel length and predictable stress, the systems are not identical. BP can have only one phonetically long vowel per word and the position of stress in a word depends on the length if its underlying vowels; in BT and S on the other hand, a word may have any number of surface long vowels, and stress is independent of vowel length, falling always on the initial syllable.

3.3 Vowel nasalization

SP has a contrast between nasalized and non-nasalized vowels not found in either S, BP or BT.[27] In BP, the SP nasalized vowels have been reanalyzed as vowel plus nasal consonant or, in some cases (notably final /ãĩ/ and final unstressed /ẽĩ/), as vowel alone:

(71) BP [pá:m] 'bread' < SP pão /pãũ/ id.
(72) BP [mú:ndu] 'world' < SP mundo /mũdu/[28] id.
(73) BP [ó:mi] 'man' < SP homem /omẽĩ/ id.

cf.

(74) BP [pá:w] 'board' < SP pau /pau/
(75) BP [mú:du] 'mute' < SP mudo /mudu/
(76) BP [fó:mi] 'hunger' < SP fome /fome/

SLCP did have nasalization earlier in its history. Dalgado reports that though nazalization had been lost in SP final /ãĩ/ and final unstressed /ẽĩ/, it had in general been retained.[29] Vasconcellos is able to confirm this from his encounter with two native speakers at the Paris world's

---

[27]Nor does BT have phonetically nasalized vowels such as those found finally in Indian Tamil (IT). Compare, for example, the following items:

BT [mǽɾæm]	IT [mérõ]	'tree'
BT [pó:ro:m ~ pó:rəm]	IT [pó:rõ:]	'we are going'
BT [pó:ra:n]	IT [pó:rã:]	'he is going'

[28]In 16th century Portuguese this may have been [mũ'ndu], i.e. /mundu/ with the earlier nasal still in place. However final nasals had certainly been lost by then.

[29]P. 11-12. Elsewhere, however, he notes: 'The tendency ... to eliminate nasalization ... is general'. (p. XXVI).

fair in 1900.[30] Among his examples are the following:

(77) SLCP sabã 'soap' (BP [səvá•m]) < SP sabão

(78) SLCP ermã 'sibling' (BP [irumá•m]) < SP irmão/irmã

(79) SLCP tẽ 'is' (BP [té:m]) < SP tem (/tẽ/)

The loss of nasalized vowels then clearly took place in Sri Lanka, and brought the SLCP vowel system closer to those of the indigenous languages.

3.4 Unrounding of /u/

Recall that in casual speech in BP /u/ unrounds to [ʉ] when unstressed, unless the preceding consonant is a labial or the vowel of the preceding syllable is round (ch. II rule (560)). This has an exact replica in BT, though here the rule operates in all but the most careful of styles.

(80) BT /iṕu/ ——> [íṕʉ] 'be'

(81) BT /itu/ ——> [íðʉ] 'this (one)'

(82) BT /a:ru/ ——> [á:rʉ] 'six; river'

(83) BT /kaLuday/ ——> [kálʉðə] 'donkey'

cf.

(84) BT /po:Tu/ ——> [po:du] 'put'

(85) BT /uTuppu/ ——> [úduppu] 'clothes'

(86) BT /pa:mpu/ ——> [pá:mbu] 'snake'

As unrounding is found in all Tamil dialects but not in Sinhala[31]

[30](1901) pp. 146-7.

[31]Neither Dalgado (1900) nor Vasconcellos (1901) mention unrounding in SLCP of Colombo in the Sinhala speaking region.

we have a clearcut case of Tamil influence on BP.

3.5 Retraction of non-low front vowels

Recall that in BP /i: i e: e/ are retracted to [ɨ: ɨ ɜ: ɜ] before rC; we also find the retraction of /i/ before rV for some speakers and that of /e/ optionally in all environments (Ch. II rule (542)). In BT /i: i e: e/ retract before retroflex consonants and /r/. In addition /e: e/ are retracted before labials and velars.[32]

(87) BT /e:ru/ ——> [ɜ:rʉ] 'climb'

(88) BT /ki:ru/ ——> [kɨ́:rʉ] 'draw'

(89) BT /verumay/ ——> [vɜ́rʉmə] 'emptiness'

(90) BT /niram/ ——> [nɨ́rəm] 'color'

(91) BT /e:Lu/ ——> [ɜ:l̤ʉ] 'seven'

(92) BT /ki:L/ ——> [kɨ́:l̤] 'down'

(93) BT /eNNay/ ——> [ɜn̤n̤ə] 'oil'

(94) BT /vi:Tu/ ——> [vɨ́:d̤ʉ] 'house'

(95) BT /iTuppu/ ——> [ɨ́d̤ʉppu] 'hip'

(96) BT /evaLavu/ ——> [ɜvəl̤əvu] 'how much'

(97) BT /e:varay/ ——> [ɜ:vərə] 'belch'

(98) BT /eppa/ ——> [ɜppə] 'when'

Cf.

(99) BT /ki:ŕay/ ——> [kí:ŕə] 'greens'

(100) BT /eŕi/ ——> [éŕi] 'burn'

(101) BT /itu/ ——> [íðʉ] 'this (one)'

---

[32] A preceding /ŕ/ may inhibit retraction. Thus: BT /ŕeNTu/ ——> [ŕén̤d̤ʉ ~ ŕɜ́n̤d̤ʉ].

(102) BT /eli/ ——> [éli] 'rat'

(103) BT /ve:cay/ ——> [vé:šə] 'whore'

(104) BT /ci:lay/ ——> [ší:lə] 'cloth'

(105) BT /enna/ ——> [énnə] 'what'

(106) BT /ivaLavu/ ——> [ívəl̤əvu] 'this much'

(107) BT /ci:kka:/ ——> [ší:kka:] 'whistling'

(108) BT /ippa/ ——> [íppə] 'now'

There is a real problem here if we want to impute the development of retraction in BP to convergence. Notice that even after discounting BP's lack of retroflexes, the BT and BP rules are not identical. We are left with pairs such as BP [sé:rə] 'wax' vs. BT [šʒ́:rɯ] 'mud' and BP [sé:vi] vs. BT [tʒ́:vi] 'goddess'. Moreover, BT lacks the the rC clusters which figure as the primary environment for the rule in BP. In what sense, then, can we say there has been Tamil influence. The similarity seems too uncanny to be from any other source,[33] but the details of the distribution do not allow us to pin down a mechanism. In fact these very details would be cited by someone arguing that the phenomena are independent.

3.6 The stops

The most salient disparity between the BP and BT consonant systems is the fact that BP has two series of stops,

[33] Sinhala has no retraction rule. E.g. [kíri] 'milk' [íḍə] permission'. As far as I can tell, the [r]'s in the three languages are the same phonetically, viz. a post alveolar tap or brief trill.

/p t c k/ and /b d j g/, while BT has only one, /p t T c k/. There is however considerable similarity on the phonetic level.

The general distribution of stops in BT is as follows: they are tense and voiceless initially,[34] finally, when geminated, and in clusters with other stops; lax and voiced following nasals; and lax, voiced and spirantized intervocalically and in clusters with liquids and glides. Each consonant has its own peculiarities, however. These are detailed in Table 9.

Assuming that these are all underlying [+tense, -voice], we would formulate three rules: (109) Lenition, which applies to all stops intervocalically and in clusters with nasals, liquids, and glides; (110) Voicing, which applies to all stops after nasals, and to all but /c/ intervocalically and in clusters with liquids and glides; and (111) Spirantization, which applies to /k c p/ and optionally to /t/ intervocalically and in intervocalic clusters with liquids and glides.

(109) BT Lenition:

$$C \longrightarrow [-\text{tense}] \;/\; [+\text{son}]\_\_[+\text{son}]$$

(110) BT Voicing:

$$\begin{bmatrix} C \\ \alpha\,\text{strid} \end{bmatrix} \longrightarrow [+\text{voice}] \;/\; \begin{bmatrix} +\text{son} \\ \beta\,\text{nas} \end{bmatrix}\_\_[+\text{son}]$$

if α is $\pm$, then β is $\pm$.

---

[34] The contrast which has developed in Indian Tamil between voiced and voiceless stops in initial position is not found in Sri Lanka Tamil. Thus IT [dó:se] but BT [tó:šə] IT [bá:ṣe] but BT [pá:šə]. For words of Dravidian origin my distributional statements are in the main true for all dialects of Tamil.

(111) BT Spirantization:

$$\begin{bmatrix} C \\ -nas \\ \alpha ant \\ -\alpha high \\ \beta cor \end{bmatrix} \longrightarrow [+cont] \; / \; [-nas]\_\_$$

optional if both $\underline{\alpha}$ and $\underline{\beta}$ are $\underline{\pm}$

We omit the further rules necessary to account for the initial variant of /c/ as well as for the various peculiarities of the spirants.[35]

Recall that in BP /b g/ and less commonly /d/ optionally spirantize intervocalically and in intervocalic clusters with /r/ (Ch. II rule (651)). The parallelism between the two Spirantization rules is evident. Vasconcellos reports that for the speakers he interviewed in 1900 "the consonants -d-, -b-, -g- between vowels are not fricatives but are pronounced as in Portuguese in initial syllables".[36] Since in Sinhala too, voiced stops do not have spirantized allophones, the development of these in BP is clearly due to Tamil influence.

We noted during the discussion of the BP rule that it applied more frequently among younger speakers. This may be taken as evidence of ongoing convergence. Finally, notice that although BP /j/ does not become spirantized, it does weaken to [y] and thus fits in with the general pattern.

---

[35] Note that (109) will assign the feature [-tense] to intervocalic /c/. Certainly lax [š] does occur in this position, though possibly it may be optionally tense. I don't know whether initial /c/ is [-tense].

[36] 1901, p. 146.

Just as BP voiced stops are similar to single BT stops in certain environments, so too there is a similarity between BP voiceless stops and BT geminates. In BP, voiceless stops have a tendency to geminate intervocalically under certain conditions, and under certain conditions BT geminates tend to degeminate. Some factors affecting these two processes are summarized in Table 10.

It can be seen immediately that these conditions operate in opposing directions with the result that quite different underlying representations yield the same phonetic output.

Unfortunately, when I was in Batticaloa I did not investigate BT degemination. However, since returning from the field, I have obtained examples of degemination in Indian Tamil and this will serve to illustrate the process. The data were collected from a single subject, who was presented with a list of words containing geminates. He was asked to repeat each word twice, (slowly and carefully and then at normal speed) and then to think of a sentence containing the word and say it slowly and at normal speed. In the examples below the stylistic variants are given in the order just described. S and F signify slow and fast respectively. The observed phonetic value of the phonemic geminates is given in brackets. Primary and secondary stress are indicated by [ˊ] and [ˋ]. The rest of the transcription is phonemic.

Table 9. Positional Variants of BT Stops

	initially etc.	geminated etc.	post-nasally	inter-vocalically
/p/	p	p	b	β
/t/	t	t	d	d~ð
/T/	ṭ	ṭ	ḍ	ḍ
/c/	š	č	ǰ	š
/k/	k	k	g	ɣ~h

Table 10. Factors Affecting BT Degemination and BP Gemination

Conditions		Effects: BT Degemination	Effects: BP Gemination
Style	careful	inhibits	promotes
	casual	promotes	inhibits
Position of stress	adjacent	inhibits	promotes
	not adjacent	promotes	inhibits
Preceding short stressed vowel		inhibits	promotes

(112) S ví:[tt]-u[kk]u

F ví:[tt]-u[k]u

house DAT 'home'

S nà:n ví:[tt]-u[k]u pò:-r -e:n

F nà:n ví:[t]-u[k]u pò:-r -e:n

I home DAT go PRES CNC

'I am going home.'

(113) S kúTu-[tt]-i[tt]u

F kúTu-[t]-i[t]u

give PAS PFC-PPL 'having given'

S nà:n inta pústaka[tt]-a | kúTu-[tt]-i[tt]u

F nà:n inta pùstaka[t]-a kúTu-[t]-i[t]u

I this book ACC give PAS PFC-PPL

va -r -e:n

va -r -e:n

come PRES CNC

'I'll give this book [to X] and come back; I'll just give this book [to X].'

(114) S cá:[pp]à:[ṭṭ]-u[kk]u

F cá:[pp]a:[ṭ]-u[k]u

food DAT 'for food'

S ní: mátya:Nam cà:[pp]a:[ṭ]u[k]u va -r -i -yá:

F ni: matya:Nam cà:[p]a:[ṭ]u[k]u va -r -i -yá:

you noon food come PRES CNC Q

'Are you coming to lunch?'

(115) S vá[ṭṭ]a[tt]-ilè:[37]

F vá[ṭṭ]a[tt]-ila

circle LOC

S inta và[ṭṭ]a[tt]-ila oru kó:Tu po:Tu

F inta và[ṭṭ]a[t]-ila oru kó:Tu po:Tu

this circle LOC a line put

'Draw a line in this circle.'

(116) S lá[čč]a[tt]i-

F là[čč]a[tt]i- pa[tt]-á:yram

lakh(100,000) ten thousand '110,000'

S èn-ki[ṭṭ]a là[čč]a[tt]i pa[tt]à:yra rú:pa: iru[k]u

F en-ki[ṭ]a là[čč]a[t]i pa[t]a:yra rú:pa: iru[k]u

me LOC lakh ten thousand rupee is

'I have a hundred and ten thousand rupees.'

The phonetic distribution of the intervocalic stops and

---

[37]Literary pronunciation.

geminate stops and their underlying sources is summarized in Table 11.

Note that because the single intervocalic stops in Tamil are voiced by an early rule, these do not fall together with the single voiceless stops produced by Degemination. I.e. in intervocalic position contrast is between single voiced stops and voiceless optionally geminate stops. The same phonetic distribution is found in BP. Sinhala can be ruled out as a model in the development of this distribution in BP, as Sinhala itself has voiced and voiceless stops, and for both of these, single stops contrast with geminates. It is interesting, therefore, that gemination was observed in Colombo SLCP by Vasconcellos and Dalgado.[38] This seems to indicate Tamil influence on SLCP as a whole (possibly via the ubiquitous Muslim traders?[39]) an interpretation reinforced by the predominance of Tamil over Sinhala loans in the various SLCP dictionaries produced in Colombo in the early nineteenth century.

It is striking that Tamil should have had such a tremendous influence on the phonetic level, but very little on the underlying representations. The only systemic effect has been the addition of low level rules to derive the new phonetic output, but this is of minor consequence as any phonetic change will of necessity entail some modification

---

[38] See note 50 p. 115.

[39] The mother tongue of virtually all Sri Lanka Muslims is Tamil.

of the P-rules. Moreover, only in the case of Spirantization can we say that a P-rule has been transferred more or less directly from BT to BP. The counterpart of Tamil Degemination is BP Gemination, a rule which has no formal similarity, and which acts in a completely opposing direction. Note that even the BT and BP Spirantization rules are not identical: they have, for example, different effects on BT /c/ and BP /j/. The generalization which this data suggests is that all direct phonological influence of one language on another takes place at the phonetic level and that systemic influences are only the secondary, induced, results of the primary phonetic influence. We shall return to this point later.

3.7 The palatals

The development of the SP palatal consonants /č ǰ l̃ ñ/ and [y][40] in BP is summarized in Table 12. It will be noticed that /ǰ l̃/, elements which do not occur in BT, have disappeared completely in the case of /l̃/, or partially in the case of /ǰ/.[41] Elements which do occur in BT (/c y ñ/)[42] are retained in BP.

The merger of /l̃ ǰ/ can also be observed in Cape Verde Creole.[43] It would appear, however, that this is an

---

[40] Treated here as an allophone of /i/ in SP.

[41] The loss of /j/ in BP is currently spreading to initial position. Cf. Ch. II rule (352).

[42] /ñ/ in BT occurs only initially and is extremely rare.

[43] B. L. da Silva.

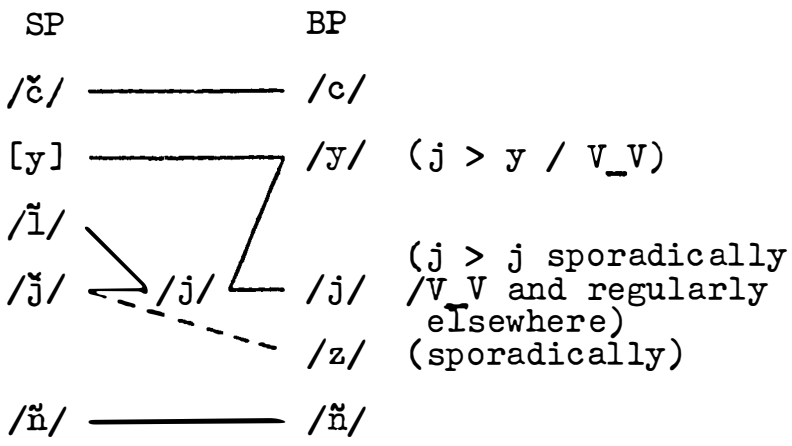
Table 11.

Intervocalic Stops and Geminates and their Sources

BT underlying	BT phonetic	BP phonetic	BP underlying
/p/	[β]	[b~β~v]	/b/
/pp/	[pp~p]	[p~pp]	/p/
/t/	[d~ð]	[d~ð]	/d/
/tt/	[tt~t]	[t~tt]	/t/
/c/	[š]	[ǰ~y]	/j/
/cc/	[čč~č]	[č~čč]	/c/
/k/	[ɣ~h]	[g~ɣ~x~h]	/g/
/kk/	[kk~k]	[k~kk]	/k/
/T/	[ḍ]		
/TT/	[ṭṭ~ṭ]		
/rr/	[ɾɾ~ɾ]*		

*[ɾ] is an alveolar stop, which contrasts with both [t] and [ṭ].

Table 12. Development of SP Palatals in BP



independent development since in the creole Portuguese of the people taken by the Dutch from South Asia to form a colony at Tugu Jakarta in the 17th century this merger does not take place. Instead /l̃/ falls together with /l/.[44] We may assume then that in Indo-Portuguese, and in particular in early SLCP, /l̃ ǰ/ were distinct. They had certainly begun to fall together in SLCP by the early 19th century, but the merger was not complete until the further change of intervocalic /j/ to /y/[45] was well under way. Below we compare evidence from the early nineteenth century dictionaries and Vasconcellos with current BP.

(117) SP filho 'son'

Fox filho

Cal fil'ho[46]

Dal filjo, filho ('more common')[47]

LV fidjo[48]

BP [fí:yu]

---

[44] Wallace 1977.

[45] Non-intervocalic /l̃/ in SP is extremely rare and no items containing it survived in BP. Thus all inherited instances of SP /l̃/ became BP /y/ (an exception is (126)).

[46] The reader is reminded that Callaway's stress indications probably have no basis in reality.

[47] j = [ǰ]

[48] dj = [ǰ]

(118) SP folha 'leaf'

Fox folha

Cal fol´ha

Dal folho

LV foya

BP [fɔ́:yə]

(119) SP mulher 'woman, wife'

Fox mulher

Cal mul´her

Dal mulhára,[49] mulher, mujá, molja

LV mudjer

BP [muyé·rə]

(120) SP alho 'garlic'

Cal ajo

Dal ajo

BP [á:y]

(121) SP olho 'eye'

Fox olho, ojo

Cal ol´ha, ojo

Dal olha, oljo, ojo ('little used')

BP [ó:y]

---

[49] Probably [mul̃ǽ·rə], with lowering of /e:/ to /æ:/ when /ə/ (<SP unstressed a) is in the next syllable. Cf. note 47 p. 112.

(122) SP olhar 'see'
Fox olha
Cal ol'ha
Dal olha, olja ('modern form')
LV oyá
BP [oyá•]

(123) SP orvalho 'dew'
Fox orevaye
Cal orval'ho
Dal or(a)vai
BP [uruvá•yə]

(124) SP embrulhar 'wrap'
Fox embria
Cal em'bria
Dal embriá
BP [embriyá•]

(125) SP trabalho 'work'
Fox travajo
Cal travalho
Dal travelho ('little used'), travajo, travai
BP [trəvá•y]

(126) SP gralha 'raven, rook'
Fox grahi
Cal grahi
BP [grá:ǰə ~ grá:yə]

(127) SP igreja 'church'
Fox greya
Cal (i)grey´a
Dal greya
BP [gré:ǰə ~ gré:yə]

(128) SP mijo 'urine'
Fox milho
mija (Fox 1820)
BP [mí:yu]

(129) SP sujo 'dirty'
Fox soujo
Cal suju
Dal sujo
BP [sú:y]

(130) SP hoje 'today'
Fox (h)ogio
Cal ho´ju[50]
Dal oljo ('modern form'), hojo ('more common'), hôjoe
LV ôi
BP [óy]

(131) SP relógio (arch. orlogio) 'clock'
Cal orlozo
Dal orlozo
BP [orló·zu]

[50]Possibly ho´jo; the printing in my copy is not clear.

Already in the dictionaries of Fox and Callaway /l̃/ > /j/ is common, and /j l̃/ > /y/ can also be found. Both authors might be suspected of scribal conservatism; the hypercorrect milho in (128) is perhaps an indication of this, though it could just as well be from a hypercorrecting informant: it is not clear from where they get their data. There certainly seems to be no aversion to listing forms with /j/ as variants of forms with /l̃/ in e.g. (121).

For most forms with historical /l̃/ Dalgado gives variants with lj (usually indicated as recent developments) and sometimes j or even y or i. Taking into account the fact that Dalgado's work is based on written sources,[51] some of which date from the beginning of the nineteenth century, we can safely conclude that by the end of the century /l̃/ had disappeared, having been replaced by either /lj/ or /j/, which had in some items further developed into /y/. As the writing of lj for SP lh could not have been modeled on SP, it probably reflected actual speech. Thus the contrast between earlier /l̃/ and /j/ was perhaps maintained in at least some items. However, the hypercorrect oljo in (130) indicates that the distinction was on its way out.

Vasconcellos reports that his two speakers did not, have /l̃/. In his scanty data SP lh shows up in some words as [ǰ], in others as [y].

In interpreting the early data from Fox, Callaway, and

---

[51]His texts stretch from 1826 to the 1890's.

Dalgado, we must keep in mind that until the middle of the nineteenth century SLCP was still widely used as a lingua franca and could therefore be expected to exhibit a certain degree of variability. Thus in the early 1800's there might have been a continuum of speakers ranging from those who consistently maintained the distinctions between /l̃ j y/ down to those who had only /y/, with mother-tongue speakers of the creole tending towards the top end of the range and non-mother-tongue speakers tending towards the bottom. Differences among regional dialects might also have been involved. Neither Sinhala nor Tamil has /l̃/. Sinhala has /j/ and /y/ and thus speakers of Sinhala would probably have tended to merge /l̃ j/ in the creole, while maintaining the distinction between these and [y]. Tamil on the other hand, has [ǰ] only as a postnasal variant of /c/, and thus Tamil speakers would probably have tended to merge creole /l̃ j/ and [y] to /y/. By the end of the nineteenth century, the language was no longer a lingua franca. Probably for most speakers /l̃/ and /j/ had merged to /j/ (possibly [ǰ ~ l̃ǰ] given Dalgado's data), and /j/ (including /j/ < /l̃/) had further developed to /y/ in some environments.

In current BP both SP /l̃/ and intervocalic /ǰ/ show up as /y/.[52] Two exceptions are (126) and (127).[53] The intermediate stage in the development /l̃/ > /j/ > /y/ is attested

---

[52]Sporadically SP /ǰ/ > BP /z/ as in (130); this appears to have been an early development.

[53]Some older speakers also have the variants [fí:ǰə ~ fí:yə] ('daughter' < SP filha).

in (126).

3.8 Retroflex consonants

BT has a retroflex series /T N L/ which BP lacks on the phonemic level. BP does, however, have retroflex [l̤ n̤] as allophones of /l n/ intervocalically and finally, following [ɔ: ɔ a: ə] (Ch. II rule (765)). As Sinhala has no retroflex liquids or nasals, this development in BP can be attributed entirely to Tamil influence.

4 Phonological Change in Generative Grammar

In this section I would like to reexamine the question of locus of phonological change - whether phonological change is grammatical or phonetic - in light of the data presented in the last section.

The generative view of historical linguistics has laid great emphasis on the claim that sound change takes place in the grammar - i.e. at some higher level of organization than the phonetic surface. One of the earliest proponents of this point of view was Halle, who for example proposed that in the study of dialect differentiation we 'focus ... on the grammars of the dialects, i.e., on the ordered set of statements that describe the data, rather than on the data directly.'[54] Kiparsky restates Halle's view that 'linguistic innovations, since they are added grammatical rules, originate in langue, in direct contradiction to the

[54] 1962 p. 342.

view which is quite explicit in Saussure and implicit in neo-grammarian and American structuralist theory that linguistic innovations originate in parole.'[55] King (1969) makes the most explicit statement of this tenet:

> ...the real question is whether sound change in the traditional sense is a proper concept at all: is it the sounds that change, or is it something else? Generative grammar maintains that it is not sounds or phones that change but grammar - a speaker's competence (Postal 1968: 269-307). Alteration in competence is reflected in performance, but not the other way around. Sound changes result from [emphasis mine] changes in competence, in the internalized system of rules for linguistic behavior. Such changes are of various kinds; rule additions and losses, reordering, simplification. These changes are not caused by tiny variations in performance that somehow seep osmotically up into competence and change it. (p. 113).

The problem with the hypothesis that grammar is the locus of all phonological change is that it is not testable on the basis of the usual type of empirical data which concerns the internal history of a given language or group of languages. Since any sound change will have grammatical correlates, supporters of the hypothesis are always able to claim that a given sound change is actually the result of a change in the grammar, and its detractors are equally free to claim the opposite. Without some kind of external evidence the whole question becomes a matter of point of view rather than a substantive issue. Situations of cross language influence often do provide external data on the motivation for a given change. Consequently, such data deserves our close attention.

---

[55] 1965 p. 1-12,

Before looking at the data, we should first ask what the standard generative model predicts about the nature of phonological influence of one language on another. Clearly if the locus of change is grammar, such influence should be discernible from a comparison of the grammars of the two languages as they change through time: we should expect that the grammar of the recipient language, in particular its phonological inventory and set of rules, should take on characteristics of the grammar of the influencing language. Let us then examine some of the phonological changes in BP brought about by indigenous influence, and see whether these predictions are born out. The first two developments considered are consistent with a grammar-centered view of phonological change. In the second two examples, however, a phonetics-oriented approach will be seen to be more fruitful.

4.1 Vowel nasalization

From a generative point of view, the loss of contrastive vowel nasality in BP is formally represented by the addition of rule (132) to its grammar.

(132) Loss of Contrastive Vowel Nasality

$$V \longrightarrow [-nas]$$

This rule addition, since it involves absolute neutralization, entails immediate restructuring: for new learners of the language there are no [+nas] vowels in the inventory and

for them (132) has the status of a redundancy rule. The influence of BT can be seen at the abstract level by the fact that the feature [nasal] has become non-distinctive in BP as it is in BT. BP has added to its grammar a redundancy rule which is identical to a BT redundancy rule. Thus, as the standard generative model predicts, the grammar of BP has moved towards that of BT.

4.2 Unrounding of /u/

The unrounding of /u/ in BP is correlated with the addition of rule (560) of Chapter II. This time there was no neutralization and thus no restructuring; the rule remains a synchronic one. Since the added rule is formally identical to the BT unrounding rule, it can again be said that the grammar of BP has become more similar to that of BT, as is consistent with the predictions of the standard model.

4.3 Retroflex consonants

The development of retroflex allophones of /n l/ in BP is another instance of rule addition (Chapter II rule (765)). As in the last example no change takes place in the underlying phonological inventory and the retroflection rule remains a synchronic rule. However, this time when we look to the grammar of BT for motivation, we find none. In BT retroflexes contrast with dentoalveolars at the underlying level, but BP has no such contrast. Furthermore, BT never had a retroflection rule which could have served as a model

for the BP rule. As a consequence the grammar of BP cannot be said to have become closer in any sense to that of BT. Hence, by simply comparing the grammars of the two languages we are unable to capture the fact that BT has influenced BP. Notice, however, that the phonetic outputs of the two grammars are more similar after the change in BP since they both now contain retroflex [n̩ l̩].

4.4 Gemination

The development of geminate allophones of voiceless stops in BP is accounted for by the addition of rule (624) of Chapter II to its grammar. Because in initial position and in clusters we find a contrast between simple voiced and voiceless stops only, these intervocalic geminates are still analyzable as underlying single voiceless stops. BT, on the other hand, has an underlying contrast between simple and geminate stops. It also has a rule which degeminates the geminates in environments complementary to those in which BP gemination occurs. BP has not incorporated geminates into its underlying representations, nor has it added to its grammar any rule modeled on a BT rule. As was the case with retroflection, the grammars of the two languages have not developed any similarity: it is only when we look at the phonetic output that we discover that BP and BT have single and geminate voiceless stops occurring in the same complementary distribution.

4.5 Implications

From the last two cases it is evident that the standard generative model of phonological change is in difficulty. By insisting on viewing all sound change as grammar change exclusively, it fails to account for processes of influence which are clearly visible in the phonetic data.

The notion of sound change as grammar change was originally invoked as an argument against the earlier concept of gradual phonetic change. Note that none of the data presented here in any way support gradual phonetic change; nor do they deny the fact that looking at the grammatical correlates of sound change often provides fruitful insights. The data do show, however, that phonological change cannot be viewed as always originating in grammatical change. We have seen that it is quite possible for grammar to change <u>as a result</u> of phonetic change. Phonological change is best seen as a two-way street combining both sound change and grammar change, though it is sometimes profitable to give more weight to one aspect rather than the other.

In cross-language influence it seems to be more often profitable to emphasize the phonetic aspects of sound change, since the development of similarity on the phonetic level does not necessarily entail the development of similarities in the grammar. This fact has led some to deny that influence has taken place if the innovations in the

recipient language can be accounted for on purely internal grounds. Welmers, for example, claims that 'language changes due to external influence are non-systematic,'[56] and argues that 'the established principles of comparative and historical linguistics and all we know about language history and language change, demand that, in divergent linguistic structures, we seek explanations first on the basis of recognized processes of internal change.'[57] Similarly, in denying any role to a (presumably Dravidian) substratum in the development of retroflex consonants in Old Indo-Aryan (OIA), Misra argues:[58]

> A sweeping generalization that "it is beyond doubt that, even where Indo-European material yields Sanskrit retroflexes, pre-Indo-Aryan and pre-Dravidian bilingualism provided the conditions which allowed pre-Indo-Aryan allophones to be redistributed as retroflex phonemes" (Emeneau 1956: 7) is unsupported by facts and is misleading. As a matter of fact, just the opposite is the case for the most part. ... In short, it is evident that the development of retroflexes in OIA is not primarily a result of non-Aryan influences but results mostly from the changes internal to the Indo-Aryan History.

Such an approach admits of inter-language influence in phonology only when borrowed lexical items are the source

---

[56] P. 3.

[57] P. 5.

[58] Pp. 64-5. Emeneau and Misra have in mind sequences of changes such as Indo-Iranian (IIr) */šdh/ > pre-OIA */ṣdh/ > OIA /ḍh/ and IIr */sdh/ > OIA /dh/, with some changes in the preceding vowel which need not concern us. This led to contrasts such as Skt. /leḍhi/ 3sg. pres. ind. 'licks' < IIr */lašdhi/ vs. Skt. /edhi/ 2sg. imper. 'be' < IIr */asdhi/ (examples from Misra p. 69).

of new contrasts - as for example with the well-known split of Old English /f/ into Middle English /f v/, occasioned by the borrowing of French words with initial [v].[59] Thus all of the instances of influence outlined here would be treated by Welmers and Misra as independent developments in BP. This is, I think, an unreasonable position, both for BP and for OIA. Given that phonetics may be the locus of cross-language phonological influence, it follows that language-internal change at the abstract level can be brought about by convergence at the phonetic level.

---

[59]Cf. Hockett p. 384; Welmers p. 2.

CONCLUSIONS

In Chapter I we traced the external history of SLCP. Briefly, Indo-Portuguese was creolized in Sri Lanka early in the 16th century and by the end of the Portuguese reign in 1658 had become a lingua franca even in areas not under Portuguese control. The Dutch continued to use SLCP as a lingua franca, and it even became their informal language of the home. Dutch was apparently used only for official purposes. During the Dutch period a cleavage developed in the creole community between the 'Dutch', who had more European blood and links to the circles of power, and the low-status 'Portuguese'. After the British took over the island in 1796, many of these 'Dutch Burghers' remained in their posts and served the new administration. The official language, however, was soon changed to English, and as the Dutch spoke SLCP at home, the Dutch language was an early casualty of the British takeover. SLCP, on the other hand, continued as a lingua franca for some time, before finally yielding to English soon after the middle of the nineteenth century. As English became the new prestige language, and as SLCP was associated with the low-status 'Portuguese Burghers', late in the nineteenth century there was a move among the Dutch Burghers of Colombo and its environs to replace the creole with English as a home language. In Batticaloa, and probably also in other areas, the Dutch and Portuguese Burghers reunited into a single community.

In Chapter II we outlined the phonology of present day Batticaloa Portuguese according to a fairly concrete generative model. We excluded from consideration words of Tamil or English origin, assuming them to be recognizably foreign. Dutch words were, however, taken to be 'native'. The major alternations were accounted for; many very rare and/or morphologized alternations were consigned to the lexicon.

In Chapter III influences of the indigenous languages of Sri Lanka on SLCP were investigated. It was demonstrated that there has been strong influence in syntax and morphology, as well as in phonology. It is not possible to attribute syntactic influence to Sinhala or Tamil alone since these languages are themselves quite similar. The major phonological influence on BP can be shown to have come from Tamil. Various traits of BP: the unrounding of /u/, retraction of non-low front vowels, the peculiar phonetic distribution of the intervocalic stops, the merger of intervocalic /j/ and /y/, and the development of retroflex [n̤ l̤], have parallels in Tamil, but not in Sinhala. Indeed, as some of these features are found in the creole of Colombo, we concluded that there must have been a Tamil influence early in the development of SLCP.

In many instances, although there were strong phonetic similarities between BT and BP, the phonological processes by which these were produced were only partially similar and sometimes, as with BP Gemination and BT Degemination,

quite dissimilar. We concluded that often the locus of influence is the phonetic rather than the abstract level; phonetic influence may induce systemic change in the recipient language, which does not lead in the direction of the donor system. In general a language incorporates the effects of phonetic influence according to the internal workings of its own phonological system and not as a function of the phonological system of the donor language. Thus phonological change in such cases is primarily phonetic and secondarily grammatical, rather than the other way around as has been claimed in recent literature.

BIBLIOGRAPHY

Abeyasinghe, Tikiri. 1966. Portuguese rule in Ceylon 1594-1612. Colombo: Lake House.

Anthonisz, R. G. 1908. The disuse of the Dutch language in Ceylon. JDBUC 1(1)29-37.

Arasaratnam, Sinnappah. 1958. Dutch power in Ceylon. Amsterdam: Djambatan.

Baldaeus, Phillipus. 1672 (1960). A true and exact description of the great island of Ceylon. The Ceylon Historical Journal 8 nos. 1-4. [a translation of the section relating to Sri Lanka of 1672. Beschrijvung der Oost Indische Kusten MALABAR en CHOROMANDEL der Zelveraangrenzende Ryken en het machtige EYLAND CEYLON ... Amsterdam.]

[Batticaloa Kachcheri Planning Unit], 1974. Statistical hand book.

Brame, Michael K. (ed.). 1972. Contributions to generative phonology. Austin: University of Texas Press.

Brohier, R. L. 1973. Discovering Ceylon. Colombo: Lake House.

Bruck, A., et al. (eds.). 1974. Papers from the parasession on natural phonology. Chicago: Chicago Linguistic Society.

Burrow, T. and M. B. Emeneau. 1961. A Dravidian etymological dictionary. London: Oxford.

Câmara, J. Mattoso Jr. 1972. The Portuguese language, tr. by Anthony J. Naro. Chicago: University of Chicago Press.

Callaway, Rev. John. 1820. A vocabulary of the Ceylon Portuguese and English languages, with a series of familiar phrases. Colombo: Wesleyan Mission Press.

_____. 1818. Vocabulary with useful phrases and familiar dialogues; in the English, Portuguese and Cingalese l ngauges. Colombo: Wesleyan Mission Press.

Carvalho, José G. Herculano de. 1961. Le vocalisme atone des parlers creoles du Cap Vert. (Actas do IX Congresso Internacional de Lingüística Românica III). Boletim de Filologia. 20:3-12.

_____. 1962. Nota sobre o vocalismo antigo português: valor dos grafemas e e o em sílaba átona. Revista Portuguesa de Filologia. 12:17-39.

Ceylon Dept of Census and Statistics. Census of population, Ceylon. 1963. Vol. 1. pt. 1.

Chomsky, Noam, and Morris Halle. 1968. The sound pattern of English. New York: Harper and Row.

Clayton, Mary L. 1976. The redundance of underlying morpheme-structure conditions. Lg. 52:295-313.

Coelho, Adolpho F. 1880-1882. Os dialectos românicos ou neo-latinos na Africa, Asia e America: Ceilão. Boletim da Sociedade de Geografica de Lisboa 2(1880):156-167; 3(1882):462-466.

Cordiner, Rev. James. 1807. A description of Ceylon. London: Longman et al.

Cumming, C. F. Gordon. 1892. Two happy years in Ceylon. Edinburgh and London: Blackwood and Sons.

Dalgado, Sebastião Rodolpho. 1900a. Dialecto indo-português de Ceylão. Lisboa: Imprensa Nacional.

_____. 1900b. Dialecto indo-português de Goa. Revista Lusitana 6: 63-84.

_____. 1903. Dialecto indo-português de Damão. Ta-ssi-yang-kuo 3:359-367; 4:515-522.

_____. 1906. Dialecto indo-português do Norte. Revista Lusitana 9: 142-166, 193-228.

Denahm, E. B. 1912. Ceylon at the census of 1911. Colombo: Government Record Office.

Dunne, P. and F. Subavakiampillai. 1950. Tamil-English dictionary. Jaffna: St. Joseph's Catholic Press.

[Dutch Burgher Union of Ceylon]. 1908. List of members. JDBUC 1(1)61-7.

Emeneau, Murray B. 1956. India as a linguistic area. Lg.32: 3-16.

Ferguson's Ceylon Directory. 1966. 108th ed; 1973. 115th ed. Colombo: Lake House.

Forbes, Major. 1841. Eleven years in Ceylon. London: Bentley.

Fox, William Buckley. 1819. A dictionary in the Ceylon-Portuguese, Singhalese and English languages...to which is prefixed a compendium of the Ceylon-Portuguese language. Colombo: Wesleyan Mission Press.

Goonetileke, H. A. I. 1970. A bibliography of Ceylon. Zug, Switzerland: Inter-Documentation Company, A.G.

Goonetilleke, William. 1892. The Ceylon Portuguese 'Diski'. Orientalist: a journal of oriental literature, arts and sciences, folklore etc. (edited by William Goonetilleke. Bombay) 4(5 and 6):89-90.

Haafner, J. 1810. (1921). Travels on foot through the island of Ceylon. London: Phillips. [original version in Dutch].

Hall, Robert A. Jr. 1966. Pidgin and creole languages. Ithaca: Cornell University Press.

Halle, Morris. 1962. Phonology in generative grammar. Word 18:54-72. Reprinted in The structure of language: readings in the philosophy of language. ed. by Jerry A. Fodor and Jerrold J. Katz. Englewood Cliffs N.J.: Prentice-Hall (1964) pp. 334-352. [Page references are from this volume.]

Hancock, Ian F. 1975. Malacca Creole Portuguese: Asian, African or European? Anthropological Linguistics 17:211-236.

Harris, James W. 1969. Spanish phonology. Research Monograph No. 54. Cambridge, Mass.: M.I.T. Press.

Hart, Thomas R. Jr. 1955. Notes on sixteenth century Portuguese pronunciation. Word 11:404-415.

Hettiaratchi, D. E. 1969. Linguistics in Ceylon, I. Linguistics in South Asia. Current Trends in Linguistics 5: 736-751.

Howard, Irwin. 1975. Can the 'elsewhere condition' get anywhere? Lg.51:109-127.

Hulugalle, H. A. J. 1969. Ceylon of the early travellers. 2nd ed. Colombo: The Wesley Press.

_____. 1973. Guide to Ceylon. 2nd ed. Colombo: Lake House.

Hyman, Larry M. 1970. How concrete is phonology? Lg.46:58-76.

_____. 1975. Phonology: theory and analysis. New York: Holt Rinehart and Winston.

ymes, Dell (ed.). 1971. Pidginization and creolization of languages. Cambridge: Cambridge University Press.

Kearney, Robert N. 1973. The politics of Ceylon (Sri Lanka). Ithaca: Cornell University Press.

King, Robert D. 1969. Historical linguistics and generative grammar. Englewood Cliffs, N.J.: Prentice-Hall.

Kiparsky, Paul. 1965. Phonological change. M.I.T. Ph.D. thesis. Published 1971 Bloomington: Indiana University Linguistics Club.

_____. 1968. How abstract is phonology? Bloomington: Indiana University Linguistics Club.

_____. 1973. Abstractness, opacity, and global rules. Bloomington: Indiana University Linguistics Club.

Knox, Robert. 1681. An historical relation of the island of Ceylon. London. Recent editions with the title An historical relation of Ceylon: 1911. Glascow: MacLehose. 1958. Ceylon Historical Journal vol. 5. 1966 Colombo: Tisara Press.

LePage, Robert B. 1959. Proceedings of the conference on creole language studies. Creole Language Studies II. London: MacMillan.

Marshal, Henry. 1846. Ceylon: A general description of the island and its inhabitants... London: William Allen and Co.

Mendis, Lionel A. 1916. Race and religion in Ceylon: a book of statistics. Dehiwala: Pearl Press.

Misra, Bal Govind. 1967. Historical phonology of modern standard Hindi: Proto-Indo European to the present. Cornell University Ph.D. Thesis.

Miller, D. Gary. 1975. On constraining global rules in phonology. Lg.51:128-132.

Pieris, Paulus, E. 1918. Ceylon and the Hollanders 1658-1796. Tellippalai: American Ceylon Mission Press.

_____. (ed.). 1929. The Dutch power in Ceylon (1602-1670). London: Curzon Press; New York: Barnes and Noble.

Postal, Paul M. 1968. Aspects of phonological theory. New York: Harper and Row.

Raghavan, M. D. 1962. Ceylon: a pictorial survey of the peoples and arts. Colombo: Gunasena.

Raven-Hart, R. undated. Germans in Dutch Ceylon. National Museums of Ceylon Translation Series I. Colombo: Colombo National Museum.

Reinecke, John E. et al. 1975. A bibliography of pidgin and creole languages. Oceanic Linguistics special publication no. 14. Honolulu: The University Press of Hawaii.

Ribeiro, João. 1925. The historic tragedy of the island of Ceilão. Colombo: Colombo Apothecaries.

Schane, S. A. 1968. French phonology and morphology. Cambridge, Mass.: M.I.T. Press

_____. 1973. Generative phonology. Englewood Cliffs, New Jersey: Prentice-Hall.

Schiffman, H. F. 1971. Reader for advanced spoken Tamil. Part II. Grammar and Glossary. U.S. Dept. of Health, Education and Welfare. Office of Education.

Schuchardt, Hugo. 1883a. Kreolische Studien. II. Ueber das Indoportugiesische von Cochim. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien, (Philosophisch-historische Klasse) 102: 799-816

_____. 1883b. Kreolische Studien. III. Ueber das Indoportugiesische von Diu. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien. (Philosophisch-historische Klasse) 103: 3-18.

_____. 1883c. Kreolische Studien. VI. Ueber das Indoportugiesische von Mangalore. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien (Philo sophisch-historische Klasse) 105: 881-904.

_____. 1889. Beiträge zur Kenntnis des kreolischen Romansich. VI. Zum Indoportugiesisch von Mahé und Cannanore. Zeitschrift für Romanische Philologie 13. 516-524.

Selkirk, Rev. James. 1844. Recollections of Ceylon... London: Hatchard and Son.

Silva, Baltasar Lopes da. 1957. O dialecto crioulo de Cabo Verde. Lisboa: Imprensa Nacional.

Silva, Chandra Richard de. 1972. The Portuguese in Ceylon 1617-1638. Colombo: Cave and Co.

Sri Lanka Dept. of Census and Statistics. 1972. Statistical pocket book of Sri Lanka (Ceylon) 1972. Colombo.

Tennent, Sir James Emerson. 1850. Christianity in Ceylon. London: Murray.

_____. 1859. Ceylon, an account of the island. 3rd ed. London: Longman et al.

Theban, Laurenţiu. 1975. From creole syntax to universal semantics. mimeo. Paper presented at the International Conference on Pidgins and Creoles, Honolulu, Hawaii, January 6-10.

Theban, Maria. 1973. Structura propoziţiei în portugheză şi indo-portugheză. Studii şi cercetări lingvistice 5: 638-45.

_____. 1974. A evolução sintáctica dos dialectos indo-portugueses. Bulletin de la Société Roumaine de Linguistique Romane 10:107-116.

Todd, Loreto. 1974. Pidgins and creoles. London: Routledge and Kegan Paul.

Toussaint, J. R. 1929. The Dutch connection with Batticaloa. JDBUC 19 (2)77-95.

_____. 1931. Batticaloa between 1766 and 1796. JDBUC 20 (3)93-106.

_____. 1942. Disuse of the Dutch language in Ceylon. JDBUC 32(92)39-47.

Valkhoff, Marius F. 1966. Studies in Portuguese and Creole. Johannesburg: Witwatersrand.

Vasconsellos, José Leite de. 1901. Esquisse d'une dialectologie portugaise. Paris: Ailland. Page references are from 2nd edition (1970) Lisboa: Centro de Estudos Philologicos.

_____. 1911. Lições de filologia Portuguesa. 4th ed. Colecção Brasileira de Filologia Portuguesa. Rio de Janeiro: Livros de Portugal.

_____. 1915. De Campolide a Melrose, relação de uma viagem de estudo. Lisboa: Imprensa Nacional.

Vos, F. H. de. 1950. Portuguese patois. JDBUC 40 (4)134-139.

Wallace, Stephen. 1977. What is a creole? The example of the Portuguese language of Tugu, Jakarta, Indonesia. Paper read at the VII Linguistic Symposium on Romance Languages, Cornell University, April 22-24.

Weinreich, Uriel. 1953. Languages in contact: findings and problems. New York: Linguistics Circle of New York, Publication no. 1. Reprinted 1974. The Hague: Mouton.

Welmers, William E. 1970. Language change and language relationships in Africa. Language Sciences 12:1-8.

Wijesekera, N. D. 1950. The people of Ceylon. Colombo: Gunasena.

Yule, Colonel Henry, and A. C. Burnell. 1886. Hobson-Jobson. A glossary of colloquial Anglo-Indian words and phrases. 2nd edition edited by William Crooke. 1903 [London?]: John Murray. Reprinted 1968. New York: Humanities Press.