Dr. P. D. Gune Memorial Lectures

INDIA AS A SOCIOLINGUISTIC AREA

P. B. PANDIT

UNIVERSITY OF POONA

PRABODH B. PANDIT (1923-)

was born in Gujarat State, India, and educated at Ahmedabad, Bombay, Paris and London. He received his Ph.D. in Indo-Aryan from the University of London in 1949. In 1955-56 he held a Rockefeller Post-doctoral Fellowship at Yale University. He taught linguistics at Gujarat University and was visiting Professor of Linguistics at the Deccan College, Poona, in 1956-57 and later on Professor at the Centre of Advanced Studies in Linguistics at Deccan College, Poona. Since 1966 he has been Professor and Chairman of the Department of Linguistics at the University of Delhi. He has also been visiting Professor of Linguistics at the University of Michigan, Ann Arbor, and Cornell University. His publications include papers on various aspects of Indo-Aryan linguistics, phonemics, historical linguistics, and socio-linguistics, which appeared mainly in *Indian Linguistics (Linguistic Society of India)* and *Language (Linguistic Society of America)*, and the books *Structure and Change in Gujarati* (in Gujarati), *Phonemic and Morphemic Frequencies in Gujarati.* ●

Dr. P. D. Gune Memorial Lectures

INDIA AS A SOCIOLINGUISTIC AREA

Dr. P. B. Pandit
M.A.Ph.D.

Price Rs 12-00

UNIVERSITY OF POOONA
GANESHKHIND
POONA 7.
1972

Printed by
M. S. Latkar,
Shreesaraswati Mudranalaya,
Plot No. 76 'E' Block,
Bhosari Industrial Area,
Poona-26.

□

Published by
Shri A. H. Thorat,
I/c Registrar,
University of Poona,
Ganeshkhind, Poona-7.

□

1100-4-1972

University of Poona,
Ganeshkhind,
Poona-7.

4th April, 1972.

PUBLISHER'S NOTE

The University of Poona has great pleasure in publishing the present lecture series given in December 1968, by Dr. P. B. Pandit, Professor and Head of the Department of Linguistics, University of Delhi. These lectures were organised by the University under the Dr. P. D. Gune Memorial Lecture Series.

This was the third series of lectures. The first one was given by Dr. Sir Ralph Lilley Turner and the second by Dr. S. M. Katre.

Students and scholars of Linguistics will find these lectures of great interest.

University of Poona
Ganeshkhind, Poona-7
5th April, 1972.

A. H. Thorat
Registrar

1. BILINGUAL'S GRAMMAR :

TAMIL–SAURASHTRI GRAMMATICAL CONVERGENCE

Professor Pandurang Damodar Gune, whose memory we respectfully evoke today, symbolises in his life and work the best traditions of scholarship in Western India, (on the fringe of the mainland that was Āryāvarta). That he died so young before he was even forty is tragic, but his recognition at such a young age marks the high point of intellectual maturity and not a scholastic respect for seniority, in our area. Secondly, by the turn of the century, our scholarship had shown a creative response to the impact of new education and scientific culture. Neither unintelligent imitation, nor isolationist resistance to the new world of learning and research could have ushered in the renaissance that was India in the late nineteenth century.

We hope and believe we are worthy successors of this tradition, worthy of the Gune memorial, a living memorial in the form of a scientific community. At least this lectureship has kept up this image where we the young and the old sit together,–I am reminded of the first series in this lectureship where along with Professor Ralph Turner I had a proud seat and role here, during his lectures–haven't you seen a photograph where young Jules Bloch is standing beside Professor Bhandarkar? Secondly by bridging the gap between generations, even in terms of the scope and theme in linguistics, by a smooth and quick progression from Dr. Gune to Sociolinguistics, we, in Poona, can claim to be equally alive and creative in our response to the expanding frontiers of science. I am honoured by this opportunity to repay my debt, may I say, to Poona.

The theme of this series of lectures is some aspects of sociolinguistics; the similarity of the title 'India as a sociolinguistic area' to the wellknown theme of 'India as a linguistic area' is not quite fortuitous. Scholars in the field of Indo–Aryan linguistics have always hinted at the limitations of the hyphenated second part, viz. Aryan; Sylvan Levi, Jean Pryzulski, Jules Bloch and

Suniti Kumar Chatterji have drawn attention to the products of language contact across linguistic-genetic-families. Emeneau's welknown article (Emeneau 1956) has focussed our attention on this bridge between the culture area and the language area. The comparativist's kit lacked the tools to take care of the cases of diffusion and convergence of linguistic features across language families; the linguistic area concept is an addition to his perspective and leads him to identify and account for more complex cases of language change. But it is much more than that. Two different languages, spoken in one culture area, may eventually develop as two styles of one language in one community, styles and varieties within one language fit different roles and functions in the grammar of culture of that community. Ferguson's notion of ' diglossia ' can be extended to cases of bilingualism where two or more languages function in one grammar of culture.

I propose to present before you two case-studies; first, a case of bilingualism, Tamil and Saurashtri bilingualism in some districts of Tamilnadu, and second, a case of speech variations in a community, in Gujarati speech community, viewed synchronically and diachronically.

Number of languages in a country is not always an index of barriers of communication in that country; a multilingual country should not be interpreted as mere juxtaposition of a number of monolingual territories. In a large multilingual country like India, continuous movement of people and ideas over a long period has resulted in a diffusion of linguistic features across language families. One of the significant features of multilingualism in India is the existence of stable bilingual or multilingual communities. People in large metropolitan centres or district towns maintain their language identity for generations, despite its ' minority ' status. They speak their own language in their domestic settings and such other dealings where the speakers of ' minority ' language come in contact with each other and they speak the majority language in other contexts. This continuous code-switching has resulted in striking similarities in the grammatical structures of these languages at colloquial level. Tamil – Saurashtri bilingualism is such a case of linguistic convergence under conditions of stable bilingualism.

Saurashtri is an immigrant Indo-Aryan language spoken in some districts of Tamilnadu. The 1961 census records 155392 Saurashtri speakers in the Madras State. The districtwise distribution of the speakers based on 1951 census data (the 1961 census has not given districtwise distribution) is as follows :

1. North Arcot	3.5 percent
2. Chingleput	2 percent
3. Salem	8 percent
4. Tiruchirapalli	8 percent
5. Tanjore	13.5 percent
6. Madurai	51 percent
7. Ramanathapuram	9.5 percent
8. Tirunelveli	2.5 percent
Others	.5 percent

The Saurashtri community consists of small islands surrounded by Tamil speakers; all the adult Saurashtri speakers are bilingual; they speak Saurashtri amongst themselves, and Tamil elsewhere. The Saurashtri speakers tend to cluster together in residential districts, they also have different customs and dress (especially women), normally they do not have marriage relations with Tamil speakers; these are some of the factors which have helped the language maintenance. The Saurashtri community is a community of weavers – especially silk weavers–and thus has a homogeneity of a guild which establishes their distinct identity in Tamilnadu.

The oral tradition regarding their separation from the parent community and the route of migration indicates three major phases in the migration of this community. Originally belonging to Saurashtra (peninsular Gujarat), they left their original home, due to the persecution following continuous attacks of the Mahmud of Ghazna (Ghazni's attacks on Somnath in 1025 A. D.) and migrated to Devagiri. Devagiri, at present a Marathi speaking area, was the capital of Yadavas and founded in 1187 A. D. In the early 14th century, it was brought under the Muslim rule of the Khiljis and Tughlaqs.

Thus, in the second phase of their migration, the Saurashtris leave Devagiri and move further south in the Vijayanagaram empire, a Telugu speaking area. The battle of Talikotta in 1565 A. D.

between the Vijayanagaram and the Muslim kingdoms could have been, once again. a reason for their further migration.

They were patronised by the Nayak King of Madurai, and they settle down in Tamilnadu around 1600 A. D. (Ramasubramanian-1968, Pp. 1-6)

Saurashtris claim that silk-weaving has been their traditional occupation from earliest times. (Grierson LSI IX. 2 447-8 has recorded the name Paṭnuli for *this* dialect; literally it means ' silk thread ' - Skt. paṭṭa + Ta. nūl). Weaving, in Gujarat, is traditionally connected with the lower castes-harijans-. During a visit to Madurai, I have noted that the Saurashtris claim a high ritual status in the caste hierarchy, and perform brahminical yajnas and I have wondered that their migration may be connected with their humble origins and search for higher status. We have no authentic data about the date of their migrations, settlement history, early documents etc. so these oral traditions have limited validity. Linguistic and ethnological work would certainly throw more light on the earlier history of the Saurashtri community.

Saurashtris do not have much literary tradition, a few works in the 19th century are cited in a bibliography published by Viprabandhu Padmanabhayyar. These are religious tracts; some manuscripts of 16 th century have been mentionēd but we have not been able to locate them. Saurashtri has no script; Tamil script, sometimes with superscript numerals over consonants to indicate voicing and aspiration, is employed to write Saurashtri. In some cases Telugu script has also been employed to write Saurashtri. There are no occasions to write or publish in Saurashtri, all Saurashtri children go to Tamil schools and for all practical purposes of written communication Tamil is used. Tamil is the dominant language. Tamil-Saurashtri bilingualism is a special case of bilingualism in which one of the languages is not spoken outside the bilingual community. There is no - or virtually no - monolingual speaker of Saurashtri. There is no norm of Saurashtri outside the Saurashtri-Tamil bilingual's speech. Saurashtri is a bilingual's mother-tongue.

The study of bilingualism takes a different turn when we move away from various taxonomies of bilingualism and pay more attention to the function of bilingualism in a given social setting. A

monolingual community which is served by one language may, under conditions of various types of language contact, come to use another language also; or, at least a part of the community may use another language. Under stable conditions of coexistence, eventually, it may happen that the functions of one language may be divided between the two languages; it may come about that in certain situations, in certain contexts, one language may be used, while in other situations and other contexts, the other may be used; each language, thus, may have specific functions. This could be described as a linguistic division of labour performed by the two languages. Observers of bilingual situations have repeatedly noted that bilingual proficiency shows noticeable variation with reference to topics and contexts. A bilingual who is an effective communicator in one language in matters of technical subjects such as law or medicine, may not be an effective communicator in the same language in a market or domestic setting; the functions of communication may not be matched for different languages. Actually it is difficult to envisage a situation in which a bilingual switches from one language to anther under exactly identical circumstances. The true bilingual; the person with a completely separate set of codes at all levels, may exist only in imagination (MacNamara 1967 p. 69, Ervin–Tripp 1967 p. 5).

This shift of emphasis from the classification of interlingual influences to the study of variations in different social contexts begins with Ferguson's notion of diglossia (Ferguson 1959). Ferguson does not directly deal with the role of two languages in a bilingual community; his frame of reference is the role of different varieties of language in a monolingual community. Ferguson has chosen four language communities, Arabic, Modern Greek, Swiss German and Haitian Creole. In each of these, two varieties exist side by side; the prestigious high variety (H) and the regional dialects-the low varieties (L). ' One of the most important features of diglossia is the specialization of function for H and L. In one set of situations only H is appropriate and in another only L, with the two sets overlapping only very slightly.

It is clear that diglossia is not limited to only these four speech communities. The Bengali diglossia with its sadhu bhasha and calit bhasha distinction, the literary and colloquial varieties of Tamil,the

Hindi-Urdu diglossia (Pillai 1959-Dimock 1959-Kelkar 1968-Gumperz-1959), the formal and informal varieties, the literary and colloquial varieties, whether it is Greek or Tamil, indicate differing social contexts, differing status of the speakers and different domains of verbal encounter. One cannot replace the other in the same act of verbal encounter.

The analogy of variety-switching is extended to code-switching by Gumperz in his studies on code switching in bilingual communities : (Gumperz 1964a, 1964b, 1966). He notes (Gumperz 1967 p. 136) :

' Although the existence of convergence can no longer be in serious doubt, linguistic analysis in bilingual communities continues to employ analytical models derived from the study of monolingual communities. Among the most frequently used measures are interference, the use of elements from one language while speaking or writing another (Mackey 1965) and contrastive analysis, in which the grammars of two languages are compared directly (Banathy, Trager and Waddle 1966). In both cases it is assumed that the structure of the languages involved is relatively uniform and is known...

If instead of starting with a priori assumption that two languages are distinct, we take the opposite view and treat them as part of a single whole, many of the difficulties cited above can be avoided. This means that in his fieldwork the linguist would disregard the speaker's view of the languages as distinct entities, and treat them as part of the same linguistic repertoire (Gumperz 1964). '

Whether it is stylistic variation among the varieties of one language or whether it is code-switching across mutually unintelligible varieties, variation is rule governed behaviour and the analyst has to bring out the complex interplay of this patterned behaviour. The models of description of monolingual communication, contrastive, interference or translatability, are not suitable because they are based on the assumption that the two languages are distinct at all levels, while in fact, convergence of the different varieties in bilingual communication has been frequently noticed; Gumperz rightly observes that language distance is not an absolute, it is a function of intensity of contact and social context (Pandit 1969 b).

We have therefore reached a stage where our techniques of analysis for diglossia or bilingualism can be fairly identical because

most bilingual situations can be diglossic. It is important to note here that the grammatical storage of all those who speak two languages need not be subjected to this analysis. Labov has drawn our attention to this possible misunderstanding : ' Gumperz suggests that we shift our view from the ' coexistent systems ' concept to that of a single repertoire which includes the entire linguistic competence of the bilingual speaker. It is important to avoid an immediate misunderstanding here. Gumperz does not suggest that the competence of every person who speaks two languages can profitably be described in terms of a single linguistic system. He refers to a bilingual community in which all the effects of social control, attrition, translation and approximation have been operative upon the systems concerned. In such a context, he argues that we may profitably view the competence of the bilingual as a single system of interrelated rules (Labov – 1969).

There are some broad similarities in the grammatical structures of Indo-Aryan and Dravidian, to which attention has been already drawn by various scholars. We have presented here a comparison of some fragments of morphology and syntax of Tamil, Saurashtri and Gujarati. The similarities between Tamil and Gujarati can be interpreted as part of the larger framework of comparison of Dravidian and Indo-Aryan. There are certain significant differences also; in as much as Saurashtri shares these differences with Tamil and not with Gujarati we are able to say that Saurashtri and Tamil grammatical structures are much more closely related to each other than either of them is to Gujarati. The constituent structures of Tamil and Saurashtri are almost identical.

The data were collected from a bilingual speaker of Saurashtri; he was asked to narrate some popular stories (mainly from the Mahābhārata) first in Saurashtri and then in Tamil. The Gujarati comparison is supplied by me.

A word about the phonology and the conventions of transcription. The ranges of variation of the allomorphs of morphemes in Saurashtri are such that substitutions of a short vowel for a long vowel or a voiceless consonant by a voiced consonant (in certain positions) is ' acceptable ' to the speaker and the listener; conventional phonemic transcription will therefore result in multiplying number of allomorphs for a morpheme. Moreover, though the speaker and

the listener perceive the phonetic differences involved, the contrasts are localised at fewer spots; and there are fewer sets of minimal or near minimal pairs in our data; this indicates that a conventional phonemic description of Saurashtri would conceal much of the relevant data which can show the selection of a variant in a given social context. Labov has suggested a more fruitful 'unit' for description of such heterogeneity in a system. Instead of choosing a point or a manner of articulation one may indicate the range of variation and indicate the context in which the particular variant is selected. The morphalternants are thus intergrated into a larger system, the social structure. A phonetic variable is 'a formal unit which represents predictable and orderly variability within the heterogeneous system' (Labov 1967 p. 3). As I have not done the fieldwork necessary for such an analysis, a listing of the major allophones of Saurashtri is given below, with a statment of distribution; some information regarding the informant's Tamil speech is presented below. I have also added a statement of the Gujarati phonemic system.

SAURASHTRI

major allophones :

Consonants :

p	b	t	d	ṭ	ḍ	c	j		k	g	
Φ	β	θ	ð	ɽ			(g^0)		(x)	h	ɦ
m		n		ɳ		ñ		ŋ			
		l		ḷ							
		r									
w		y									

Stops

In the initial position :

(i) ṭ ḍ do not occur

(ii) in a number of lexical items, continuous variation (free variation) between the voiceless and voiced varieties, is noticed while in a number of other lexical items, no such variation is noticed.

(iii) all the stops other than ṭ ḍ and c can occur aspirated or unaspirated, but in a number of items the aspirated variety varies with the unaspirated variety (but not vice–versa).

In the medial position:

(i) there is continuous variation in a number of items between the voiceless and voiced variation while in a number of items no such variation is noticed.

(ii) the stops p b, t d, ṭ ḍ and k g respectively vary freely with the respective spirants Φ β, θ ð, ɾ (flap), x g^0 h. Thus, in a number of items the range of variation could be k – g – x – g^0 – h.

(iii) aspirated stops do not occur.

Final position:

(i) c, j do not occur.

(ii) aspirated stops do not occur.

(iii) there is free variation between voiceless and voiced consonants.

Nasals: –ṉ occurs only medially.
–ŋ occurs only before their homorganic consonants.

Laterals: –ḷ occurs only medially.

Vowels:

i	u	ii	uu	ẽ	õ
e	o	ee	oo	ãã	
a		aa			

The short and long vowels are partially complementary in their distribution: the vowels in monosyllabic words are predictably long, and those that occur in the final position of polysyllabic words are predictably short. In other positions, e.g. initially and medially in polysyllabic words, length contrast is noticeable. The incidence of 'free variation' in the vowel system is less than the incidence of free variation in the consonant system. Nasalized vowels occur only medially.

Tamil :

The Tamil of the Tamil-Saurashtri bilingual shows much more consistency *compared* to his Saurashtri. The incidence of free variation between phones is much less.

Consonants -

p b	t d	ṭ ḍ	c j	k g
f	s			
m	n n̲	ṇ	ñ	ŋ
	l	ḷ		
	r	ṛ		
v			y	

f and p vary in English loan-words, e.g. kaapi∼kaafi. Geminate stops and affricates occur intervocalically and are in complementary distribution with non-geminated stops elsewhere; in a few cases, s occurs initially where it varies freely with c. ṛ and ḍ are in complementary distribution : ṛ occurs only intervocalically. l and ḷ vary freely medially in a few cases. Velar and palatal nasals ŋ and ñ occur only before homorganic consonants; dental nasal n̲ occurs before d; alevolar nasal n occurs elsewhere and sometimes it varies with retroflex nasal ṇ in some cases.

Vowels :

	i̇	u	ii	uu	ɛ	ãã	ĩ	ɔ̃
e	ə	o	eeɛɛ	oo				
ɛ	a	ɔ		aa				

(i) initial i, ii, and o, oo have a y and w colour, elsewhere simple i, ii and o, oo occur.

(ii) ɛ, ɛɛ, and ɔ, ɔɔ occur in word final position and when the following vowel is central low (a) in the next syllable. They are in complementary distribution with e, ee and o, oo respectively.

(iii) u and i̇ (unrounded back vowel) are in complementation: u occurs in the initial syllable, i̇ elsewhere.

(iv) nasalised vowels occur only finally.

The vowel system therefore can be stated as follows :

short vowels :

i u

e a o

long vowels :

i: u:

e: a: o:

nasalised vowels :

ã ãã ĩ õ

Gujarati phonemic system :

p b	t d	ṭ ḍ	c j	k g	
	s	(z)	š		h
	l	ḷ			
m	n	ṇ		N	
v			y		

(i) [ḍ] and [ṛ] are in complementary distribution.

(ii) [ṇ] and [ḷ] occur non–initially.

(iii) aspirated and unsapirated consonants contrast (they can be interpreted as consonant + h).

(iv) velar nasal ŋ, palatal nasal ñ and nasalied vowels are in complementary distribution and can be grouped under one N phoneme(transcribed as ~).

Vowel phonemes :

i		u
e	ə	o
ɛ	a	ɔ

i) short and long vowels are in complementary distribution.

(ii) contrast between e, ɛ and o, ɔ is neutralised in final position in polysyllabic words and when they are nasalised.

The major sentence type in all these languages is composed of subject which consists of a noun phrase and a predicate which

consists of a verb phrase, the latter in this may be further decomposed into noun phrases and the verb which may be either a simple verb or a verb and an auxiliary. We will refer to this as clause type I. There is another equally common type in which there is no verb in the predicate, we will refer to this as clause type II.

Clause type I.

1. 'Ram sleeps'

1a. G. Ram sue chɛ
1b. S. Ramən ninjərəs
1c. T. Ramən tu:mguhiran

G sue chɛ 'sleeps' consists of the present indicative construction in which the verb (su-) and the auxiliary (ch-) are inflected for person and number su- e ch-ɛ. S and T do not have such a construction; S and T agree in having a verb inflected for the present indicative third-person singular.

2. 'Vikram praised Govind'

2a. G. Vikram-e Govind-ne vəkhaṇyo
2b. S. Wikrəmən Go:windən -uk mecces
2c. T. Wikrəmən Go:windən -əy meccinan

Note that while the subject vikrəm- in Gujarati is followed by the (*ergative*) *agentive suffix* -e, the subjects in ST are in nominative. G construction is ergative. The object Gowind-takes the respective accusative suffixes in all the three languages.

In the following two examples, the subject is in the nominative is ST as in the 2 above, while in G it is *agentive*; the subject could be in nominative case in G also, but this would be determined by the selection of the verb. The use of ergative is determined by the choice of the verb.

3. 'Vikram invaded (on) our country'

3a. G Vikrəm-e əmara gam upər humlo karyo
or G Vikrəm əmara gam upər cəḍi avyo
3b. S Wikrəmən əwre gam-u-hollo pəḍe khədes
3c. T Wikrəmmə engəḷnaṭṭ-in mi:du pəḍəy eḍuttan.

4. 'Ram passed (in) the examination'
4a. G Ram-e pərikša pas kəri
or Ram pərikša mã pas thəyo
4b. S Ramən pəricce-m pas keres
4c. T Ramən pəriccəy-il pas ceydan.

Note that the object NP in all the languages is followed by the same postposition viz meaning 'on' and 'in' in 3 and 4 respectively.

5. 'Ram told a story to Vikram'
5a. G Ram-e vikrəm-ne ek vat kəhi
5b. S Ramən oṇṭe kheni wikrəmən jo:ḷ sənges
5c. T Ramən oru kədəy wikrəmən iḍəm connan

6. 'Ram wrote a letter to Vikram'
6a. G Ram-e vikrəm-ne kagəḷ ləkhyo
6b. S Ramən oṇṭe kəḷjəm wikrmən-uk likkes
6c. T Ramən wikrəmən-ukku oru kəḍidəm eḷudinan

Note the ergative in G and also the preferred order in G in which the direct object immediately precedes the verb, while in ST the preferred order is sub. + dir. obj. + indir. obj. + verb.

7. 'This story is funny'
7a. G a vat rəmuji chɛ
7b. S elle kheni weḍik-kən se
7c. T ində kədəy weḍikkəy ahə irukkirədu

8. 'Govind is a manager'
8a. G Govind mɛnɛjər chɛ
8b. S Gowindən oṇṭe moduraḷi kən se
8c. T Go:windən oru modəlaḷi ahə irukkiran

9. 'Ram is well'
9a. G Ram ṭhik chɛ
9b. S Ramən səwkyəm kən se.
9c. T Ramən sawkyəm ahə irukkiran

In 7, 8 and 9 with the copula 'to be', the predicative attribute 'funny' in 7, the predicative nominal 'manager' in 8 and the predicative adverbial 'well' in 9 are followed by adverbial formatives (verbalisers) in S and T; literally translated, the predicative attribute in 7 would be 'as fun', the predicative nominal in 8 would be 'as a

manager', and the predicative adverbial in 9 would be 'as well'. Gujarati does not have such a construction.

10. 'Ram can do this work'
10a. G Ram a kam kəri šəke chɛ
10b. S Ramən *or* Ramən-hal elle kam ker-ette musəy
10c. T Ramən *or* Ramən-al ində kariyəttay cey-yə muḍiyum

The verb phrase in the above sentence consists of a modal- 'to be able' šək- in Gujarati, mus- in Saurashtri and muḍi- in Tamil. The modal verb phrase in all the three GST above consists of a verbal base with infinitive marker + modal verb to which person-number and tense (if any) suffixes are attached. Optionally the subject in ST is in instrumental- in Gujarati the subject is in the nominative. (Moreover, there are restrictions regarding the tense suffixes that can be affixed to the modal base-these restrictions are the same for Saurashtri and Tamil).

Clause Type II.

The copula is deleted in this type; these sentences are also called equational sentences. The copula is deleted in ST but not in G (except in special cases of emphasis or style). The deletion in ST takes place only if the deleted copula is in the present tense and the predicate phrase consists of a noun phrase which may be either an adjective + noun or a noun; if the predicate phrase consists of an adjective or an adjective phrase then the copula is not deleted in Saurashtri and it is optionally deleted in Tamil. The verbalizers which are suffixed to the adjectives are also deleted when the copula is deleted.

11. 'It is a book'
11a. G e cɔpḍi chɛ
11b. S telle boggu.
11c. T ədu puttəhəm

12. 'Ram is a boy'
12a. G Ram chokro chɛ
12b. S Ramən oṇṭe beṛko
12c. T Ramən oru pəyyən

13. 'Ram is a good man'
13a. G Ram sarɔ maṇes chɛ
13b. S Ramən cokkəṛ menik
13c. T Ramən nəllə mənidən

14. 'They are good boys'
14a. G e sara chokra–o chɛ
14b. S tenu cokkəṛ beṛkan.
14c. T əwərhəḷ nəllə pəyyəngəḷ

15. 'This is fresh milk'
15a G a tajũ dudh chɛ
15b. S ye cokkəṛ du:t
15c. T idu pudusa:na pa:lu

16. 'This is a clean shirt'
16a. G a saph khəmis chɛ
16b. S ye oṇṭe suddam cokkay
16c. T idu oru suddama:na satte

17. 'This milk is fresh'
17a. G a dudh tajũ chɛ
17b. S ye du:t cokkəṛ ga se
17c. T inde pa:lu pudusu (a: irukku)

18. 'This shirt is clean'
18a. G a khəmis saph chɛ
18b. S ye cokka:y suddam ga se
18c. T inde saṭṭe suddam (a: irukku)

Note that there is no gender–number agreement between the adjective and substantive in ST, Gujarati has a gender–number agreement between the two. It is possible to consider the copulaless sentences in the clause type II having an underlying structure with a copula, but, the comparisons offered here are of the surface structure.

Passive.

The process of transforming active to passive is identical in ST, viz. The original nominative subject becomes the agentive, the original object becomes the new subject and the original main verb stem becomes the infinitive and the tense, gender, number and

person inflections are suffixed to the passive auxiliary which follows the infinitive main verb. The passive auxiliary stem in Saurashtri is po:d and the Tamil stem is paḍu. While the transposition and transformation of the subject-object in the change from active to passive is shared by Gujarati, it has no passive auxiliary; instead, a passive suffix (-a-) is attached to the verb stem in Gujarati. Moreover, in the ergative constructions in Gujarati where the subject is already in the instrumental it is replaced by the subject with instrumental -thi which co-occurs with the passive stems in Gujarati.

The following sentences are the passive forms of sentence 2:

19. 'Govind was praised by Vikram'
19a. G Govind vikrəm-thi vəkhaṇ-a-yo
19b. S Go:windən wikrəmən-hal məcci-ni poḍes
19c. T Go:windən wikrəmen-al-məcc-ə pəṭṭan

When the transitive verbs are of the 'give-ask' type which take direct and indirect objects, the passives can be of two types in ST.

20. 'Ram gave a fruit to Govind'
20a. G Ram-e Govind-ne phaḷ apyũ
20b. S Ramən Go:windən-uk oṇṭe poḷḷo diyes
20c. T Ramən Go:windən-ukku oru pəḷəm koḍuttan

The passive forms are as follows :

21. 'Govind was given a fruit by Ram'
21a. G Same as 20a G
21b. S *Go:windən* Ramən-hal oṇṭe poḷḷo de-ni *poḍes*
21c. T *Go:windən* Ramən-al oru pəḷəm koḍu-kkə-*pəṭṭan*

22. 'A fruit was given to Govind by Ram'
22a. G *ek phəl* Ram-thi Govindne *əpayũ*
22b. S oṇṭe poḷḷo *Gowindən-uk* Ramən-hal de-ni *poḍes*
22c. T oru *pəḷəm Go:wində-ukku* Ramən-al koḍu-kkə *pəṭṭədu*

Nominalisation.

Relativisation : ST do not have relative pronouns. The relative function is performed as follows : endocentric construction

is formed by shifting one of the constituents of the clause to the head position and making it a nominal and by suffixing the relative participle marker to the verb stem to form the attributive. The verb in the participial phrase obligatorily occurs in the final position and the resultant attributive string always precedes the head.

Gujarati has a relative pronoun je and the correlatives te and e (which are anaphoric) but the relativisation is optionally permitted without the use of these pronouns; this relativised construction is structurally comparable to ST relativised construction, with the difference that when the verb is transitive, the subject of the embedded sentence is in the agentive case in the relativised phrase (when it is not already in instrumental case – in ergative). When the verb is intransitive the use of relative pronoun is obligatory in G.

23. 'The boy laughed'
23a. G chokro həsyo
23b. S elle beṛko əses
23c. T ində pəyyən cirittan

24. 'The boy who laughed ...'
24a. G *je* chokro həsyo
24b. S əse beṛko
24c. T cirittə pəyyən

25. 'Ram saw the boy'
25a. G Ram–e chokro joyo
25b. S Ramən elle beṛkak siyes
25c. T Ramən ində pəyyənəy parttan

26. 'The boy whom Ram saw'
26a. G Rame joy–elo chokro *or* je chokro Ram–e joyo
26b. S Ramən siye beṛko
26c. T Ramən parttə pəyyən

27. 'Ram gave a book to the boy'
27a. G Ram–e chokra–ne cəpdi api
27b. S Ramən elle beṛk–ak (telle) boggu diyes
27c. T Ramən ində pəyyənukku (əndə) pustəhəm koḍuttan

28. 'The boy to whom Ram gave the book'
28a. G *je* chokra–ne Ram–e cəpdi api
28b. S Ramən boggu *diye* beṛko
28c. T Ramən pustəhəm *koḍuttə* pəyyən

Note that the determiners (demonstrative) elle/ində in ST (25b, 25c) are dropped when the relative participle is suffixed to the verb (26b, 26c).

There are further restrictions on the optional absence of the relative je in G; note that it cannot be optionally dropped in 28 even though the verb is transitive. The rules for relativisation are identical in ST. When the constituent shifted to the head position is adverbial — time, location — the relativisation rules are the same as in above examples.

29. 'Ram came on this day'
29a. G Ram a divse avyo
29b. S Ram elle–dinnu awes
29c. T Ramən ində dinəm wəndam

30. 'The day on which Ram came'
30a. G je divse Ram avyo
30b. S Ramən awe dinnu
30c. T Ramən wəndə dinəm

31. 'Ram lived at this place'
31a. G Ram a jəgae rəhyo
31b. S Ramən elle thamum jiwes
31c. T Ramən ində eḍəttil wəcittan

32. 'The place where Ram lived'
32a. G je jəgae Ram rəhyo
32b. S Ramən jiwe tham
32c. T Ramən wəcittə eḍəm

Even though the clauses belonging to the Type II in ST occur without a verb (copula), when they are nominalised, they mostly take the verbs 'be' S. se– T. iru and 'become' S. ho– T. ahu, which occur finally in the relativised phrase (The ST verbalisers are teno/wən in 34b and 34c. (cf. 17 and 18).

33. 'This boy is good'
33a. G a chokro saro che
33b. S elle beṛko cokkəṛ teno
33c. T ində peyyən nəlləwən

34. 'The boy who is good'
34a. G je chokro saro chs
34b. S cokkəṛ teno hoye beṛko
34c. T nəlləwən anə pəyyən

35. 'This man has (that) disease'
35a. G a maṇəs ne (pelo) vyadhi chɛ
35b. S elle menikuk (telle) wyadi
35c. T ində mənidənukku (əndə) wyadi

36. 'The man who has the disease'
36a. G je maṇəs ne vyadhi chɛ
36b. S wyadi sette menik
36c. T wyadi uḷḷə mənidən

37. 'The disease which the man has'
37a. G a maṇəs ne je vyadhi chɛ
37b. S elle manikuk sette wyadi
37c. T ində mənidənukku uḷḷə wyadi

38. 'This man's disease'
38a. G a maṇəs no vyadhi
38b. S elle meniku wyadi
38c. T ində mənidənuḍəyə wyadi

Factive nominalisation – Single or complex sentences can be converted into factive nominals in ST by adding the infinitive of the verb 'to mean' to the simple or complex sentence. The nominalised string will have the structure

sentence + verb 'to mean' in infinitive

In G the nominalised string consists of the sentence followed by a correlative e and the possessive marker -n-. Alternatively, the nominalised string would have the verb in the infinitive and the subject will be followed by the possessive marker (e. g. Ram came-Ram's coming-G. Ram avyo – Ram nū avvū) and the nominalised string will be followed by the possessive marker Ram–*na* avva–*no* – 'of Ram's coming'.

39. 'The fact that Ram will come here gave me pleasure'
39a G Ram əhī avše *e no* məne sentoš thəyo
39b S Ramən e:t əwəy *menette* mogo sənto:s diyes
39c T Ramən imgu wəruwan *enbadu* enəkku səndo:ṣəm koḍuttədu

40. 'Everyone knows that Ram will come here'
40a G badha — Ram əhī avše – *e* jaṇe chɛ
40b S əski tenko – Ramən e:t əwəy *nenette* kəlay
40c T ello:rukkum – Ramən imgu wəruwan–*enbədu* teriyum

3. Clause nominalisation in ST can also be established into two similar types viz. those in which an infinitive is used at the end of the nominalised string and those in which the main verb is converted to a noun head forming a genitival attributive phrase. G nominalisations are also generally similar to the above; the difference is the subject (noun or pronoun) of the sentence which is being nominalised is also an attributive being followed by the possessive marker –n– (see 38).

41. 'You are in such a state of drunkenness'
41a G tū awi nəšavaḷi sthiti mā chɛ
41b S tu: iso keḷḷu bo:dam rhariyo
41c T ni: ippədi kəḷḷu bo:dəyil irukkiray

42. 'It is wrong for you to be in such a state of drunkenness'
42a G ta–rū avi nəšavaḷi sthitimā hovū khərab chɛ
42b S tu: iso keḷḷu bo:dam rhətte təppu
42c T ni : ippədi kəḷḷu bo:dəyil iruppədu təppu

43. 'you play — your playing'
43a G tū khele chɛ — ta–ro khel
43b S tu kheleriyo — tore khel
43c T ni wiḷəyaḍuhiray — unnuḍəyə wiḷəyaṭṭu

The example in 43 is restricted in Saurashtri, very few verbs (mainly affirmative intransitive verb phrases) participate in this construction while it is common in G and T

Verb morphology.

The grammatical categories of person, number, tense, and affirmation and negation in ST have some significant similarities vis á vis Gujarati. It will also be seen that the dissimilarities in ST are also equally noteworthy especially because the grammatical overlap between the two languages is least in this area. One can notice that the similarities are more significant than the dissimilarities.

The person-number sets of Gujarati verb paradigms do not recognize the human-nonhuman categories (or the rational-irrational categories). While Tamil and Saurashtri have a human-nonhuman category : the human-nonhuman forms contrast in sg. in both the languages. Tamil has further distinction of m. and f. in the human category. Tamil has distinct forms for all the three persons in the non-future tenses while Saurashtri has only the first person and non-first person distinction in future and no person distinciton in the non-future tenses. Gujarati has three person distinctions in the future and the present tenses. On the whole, the model for Saurashtri is Tamil, with lesser differentiation than Tamil.

ST have negative verb forms while the verb in Gujarati is followed or preceded by negative particles or adverbials. Saurashtri has far less differentiation than Tamil in the negative verbs also.

The following tables give a partial account of the verbs in ST and G.

FUTURE—AFFIRMATIVE

	Saurashtri	*Tamil*
Person	I and non-I	I, II, III
Number	sg. vs. pl. except in non-human	sg. vs. pl. except in non-human
Gender	non-distinctive	distinctive in III human
Main verb structure	verb stem + inflection	verb stem + tense + inflection

GUJARATI

Person	I II and III
Number	sg. vs. pl. except in III
Gender	non-distinctive
Main verb structure	verb stem + tense + inflection

TAMIL

uṭkaru ' to sit ' Future affirmative

	Sg.	*Pl.*
1st	uṭkaruwe:m	utkaruwo:m
2nd	uṭkaruway	uṭkaruwi:rhəḷ
3rd H m.	uṭkaruwan	
H f.	uṭkaruwaḷ	uṭkaruwarhəḷ

SAURASHTRI

bi:s 'to sit' Future affirmative

	Sg.	*Pl.*
1st	bisu	bisuwe
non-1st H	bisəy	bisən
non H	bisəy	

GUJARATI

	Sg.	*Pl.*
1st	bɛsiš	bɛsišũ
2nd	bɛsiš	bɛsišũ
3rd	bɛšše	

The negative verb forms for future :

TAMIL

1st	uṭkarəmaṭṭen	uṭkarəmaṭṭo:m
2nd	uṭkarəmaṭṭay	uṭkarəmaṭṭi:rhəḷ
3rd H m.	uṭkarəmaṭṭan	
H f.	uṭkarəmaṭṭaḷ	uṭkarəmaṭṭərhəḷ
non H	uṭkaradu	

SAURASHTRI

	Sg.	*Pl.*
All persons	bisna	bisnan
non H	bisna	

Gujarati does not have a negative future paradigm but the future is negativised by a negative particle nəhi (for all persons and numbers) which follows the future verb.

PRESENT AND PAST

	Saurashtri	*Tamil*
Person	non-distinctive	I, II and III distinctive
Number	Sg. and Pl. except in non-human	Sg. and Pl.
Gender	m. f. sg. distinctive	m. f. sg. distinctive
Main verb structure	verb stem + tense + inflection	Same

Gujarati – Present

Person	I, II, III distinctive
Number	Sg. and Pl. in I and II, distinctive
Gender	non–distinctive
Main verb structure	verb stem + inflection

The past in Gujarati is a participle inflected like the Gujarati nouns – for three genders and two numbers. The main verb structure is verb stem + tense + noun inflection.

Tamil – Present Affirmative

	Sg.	*Pl.*
1st	uṭkaruhiren	uṭkaruhiro:m
2nd	uṭkaruhiray	uṭkaruhi:rhəḷ
3rd H m.	uṭkaruhiran	
f.	uṭkaruhiraḷ	uṭkaruhirarhəḷ
non H	uṭkaruhirədu	uṭkaruhīdrənə

Saurashtri – Present Affirmative

	Sg.	*Pl.*
All persons		
m.	bisəres	
H		bisəryas
t.	bisəris	
nonH	bisəres	

Gujarati – Present

	Sg.	*Pl.*
1st	bɛsũ	bɛsie
2nd	bɛse	bɛso
3rd	bɛse	

The present negative in Tamil is one form for all persons, numbers and genders: uṭkarəwilləy. In Saurashtri the negative is formed by replacing –s by –ni : bisəreni, bisərini, bisəryani, and in Gujarati a negative particle nə precedes all the verb forms or nəhi (more emphatic negative) follows all the verb forms.

Tamil – Past Affirmative

	Sg.	*Pl.*
1st	uṭkarnde:n	uṭkarndo:m
2nd	uṭkarnday	uṭkarndi:rhəḷ
3rd H m.	uṭkarndan	
f.	uṭkarndal	uṭkarndarhəḷ
nonH	uṭkarndadu	uṭkarndənə

Saurashtri–Past Affirmative

All persons	*Sg.*	*Pl.*
H m.	bises	
f.	bisis	bisyas
non H	bises	

The negative past form in Tamil is uṭkarwilləy (same as the negative present form for all numbers, persons and genders. The Saurashtri negative forms are formed by replacing –s by –ni:

m. sg. biseni f. sg. bisini pl. bisyani

The Gujarati negative forms are preceded by nə or followed by nəhī (as in the present).

The patterning is similar in Saurashtri and Tamil: both have the same categories; the present and past sets in both are identical. The negative forms are also structurally more similar: ST have a negative suffixed to the verb, and the person–number distinctions are not maintained in the negative.

Various modals and compound verbs, while comparable to each other in all the three languages, are closer to each other in TS than either of them to Gujarati. Thus, in the imperative, all the three maintain the sg–pl distinction (in the II person), the negative marker always comes after the verb stem (in Tamil, it is followed by the plural marker) while in Gujarati or Marathi the negative (nə or nəko) can precede the verb stem. (It may also be noted that the negative in Saurashtri in the imperative and precative is nəko. There may be more cases of Marathi items in Saurashtri).

Negation.

The negation in the clause type II is similar in Tamil and Saurashtri. The negative particle comes at the end of the sentence.

Gujarati has a negative auxiliary verb nəthi which occurs after the non-future verbs, while the negative particle nə occurs before the verb and nəhi after the verb. The constituent negation which involves a totality determiner in the affirmation also follows the same pattern.

44. 'Govind is not your friend'
44a G Govind taro mitrə *nəthi*
44b S Go:windən togo singeyd nha
44c T Go:windən unəkko nəṇbən illəy

45. 'Ram is not well'
45a G Ram saro net'hi (e. g. Ramne ṭhik nəthi)
45b S Ramən səwkyəm nhi
45c T Ramən səwkyəm illəy

46. 'Ram never laughs' (aff. Ram always laughs)
46a G Ram kyarey həsto nəthi
46b S Ramən khobbinu əsəriyo nha
46c T Ramən oru poḻudum ci:rikkirədilləy

47. 'No boy is good' (aff. all boys are good)
47a G koy chokro saro nəthi
47b S konni beṛko cokkəṛ teno nha
47c T endə (or, oru) pəyyənum həlləwən illəy

This fragmentary comparison of T S and G may be considered as a prima facie case to interpret the grammar of T and S as one set of rules where only the low level rules such as selection of the lexical items being special for T or S.

The convergence of constituent structures and grammatical categories in Tamil and Saurashtri cannot be a result of a linear word to word replacement. One could produce a literal translation from one to the other and still retain different grammatical structures. The 'bilingual' speaker in the TS situation interprets the two codes as one and operates with the same set of rules. The choices before him are set of TS morphemes and constructions. The syntax rules are largely the same, with greater differentiation at the lower level rules.

2. BILINGUAL'S GRAMMAR :

GUJARATI–TAMIL–SAURASHTRI NUMBER NAMES*

The set of number names in a language can be easily detached and analysed independently. The grammar of number names of the three languages Gujarati, Tamil and Saurashtri is presented here. The Gujarati and Tamil grammars are presented to bring out the general similarities and differences between Indo Aryan and Dravidian number systems. The base rules for both are almost identical; but Tamil and Saurashtri number systems are more identical; they can be generated by the same model, though the terminal vocabulary in both is different. The rules for the grammar of Gujarati number names are presented in details and they are followed by Tamil and Saurashtri rules.

The set of number names consists of a limited lexicon and a limited number of rules which produce the infinitely denumerable set of number names. The rules must produce only the acceptable number names, and they should also provide for possible synonymous expressions of the number names, e.g. the following three are synonymous in Gujarati ** :

pənder sɔ	' fifteen hundred '
dɔḍh həjar	' one and a half thousand '
ek həjar pāc sɔ	' one thousand and five hundred '

---

* I am indebted to Dr. Austin Hale of Summer Institute of Linguistics (U. S. A.) for reading an earlier draft of this section and making valuable suggestions.

** Gujarati and Tamil are presented in conventional transcription. Saurashtri has also been transcribed in conventional phonemic transcription; but it should be noted that the billingual's phonology presents some problems e.g. minimal pairs are few, incidence of 'free variation' is much greater in Saurashtri e.g. medially, voiceless and voiced stops vary freely in many cases; so does vowel length. These 'free variations' may not be 'free' if viewed in a sociolinguistic framework. Here, we have transcribed all data in conventional phonemic transcription.

The above number (1500) can also be stated by merely reproducing the digits in the order in which they are written –thus : ek – pāc – mīḍũ – mīḍũ – (one – five – zero – zero –); but this is not a number name, it is a dictation label for number names. We have therefore not taken into account this type of number description.

For all practical purposes the highest number name unit in Gujarati is əbəj i.e. 1000000000. Units above this can be made available from any Sanskrit lexicon, but they are never used; we have therefore stopped at əbəj in our lexicon.

The current conventions in generative grammar regarding the presentation of rules have been observed. Thus, ⟶, ( ) and { } respectively indicate 'rewrite', optional inclusion and obligatory choice of the included elements; : is used to indicate introduction of terminal symbols. ⟶ is used for T rule (transformation rule) as well as for morphophonemic rule; as we have only one T rule in this grammar, there is no chance of confusion. I have followed A. Van Katwijk's model ' A Grammar of Dutch Number Names ' (Foundations of Language, 1. 1. 1965) with suitable alterations for describing the Gujarati number names.

GUJARATI

Non-terminal vocabulary : R_0 R_1 R_2 R_3 R_4
D^1 D^2 D^3 D^4
Ψ^1 Ψ^2 Ψ^3 Ψ^4
I XI X XX C

Terminal vocabulary : mīḍũ (sun)
ek, bɛ, trəṇ, car, pāc, chə,
sat, aṭh, nəv (1 – 9)
dəs, əgyar, bar, ter, cəvd,
pəndər, soḷ, səttər,
əḍhar, ogṇis (10 – 19)
vis (20)
sɔ (100)
həjar (1000)
lakh (100000)
kəroḍ (10000000)
əbəj (1000000000)
pa (1/4), ərdho (1/2), pɔṇo (3/4).

PRODUCTION RULES :

1. Integer $\longrightarrow$ $\left\{\begin{array}{l} \text{zero} \\ \Psi_4\ (\ [\Psi_2 \atop \Psi_1]\ ) \\ \Psi_2 \quad (\ \Psi_4\ )\Psi_3 \\ R_i \end{array}\right\}$

Of the four possible applications of this rewrite rule, the first gives us zero (mīḍū – sun); the second and third supply us with the alternative synonymous expressions referred to above; these three will be taken up later. The fourth, R_i can be expanded as follows :

2. $R_i \longrightarrow (\ (\ \Psi_4\ )\ \Psi_3\ D_i\ )\ Ri - 1$

Rule 2 is recursive so that R_i : ($R_0\ R_1\ R_2\ R_3$ and R_4) can be introduced again on the left, e.g.,

$R_4 \longrightarrow (\ (\ \Psi_4\ )\quad \Psi_3\quad D^4\ )\quad R_3$
$R_3 \longrightarrow (\ (\ \Psi_4\ )\quad \Psi_3\quad D^3\ )\quad R_2$
$R_2 \longrightarrow (\ (\ \Psi_4\ )\quad \Psi_3\quad D^2\ )\quad R_1$
$R_1 \longrightarrow (\ (\ \Psi_4\ )\quad \Psi_3\quad D^1\ )\quad R_0$

3. $R_0 \longrightarrow \Psi_1\ (\ \Psi_4\ )\ \Psi_3$

4. $\Psi_1 \longrightarrow (IC)$

5. $\Psi_2 \longrightarrow \left\{\begin{array}{l} XI \\ XX+1 \end{array}\right\}\ C$

6. $\Psi_3 \longrightarrow (\ \left\{\begin{array}{l} I \\ X \\ XI \\ XX(I) \end{array}\right\}\ )$

7. $\Psi_4 \longrightarrow (\ \left\{\begin{array}{l} \frac{1}{4} \\ \frac{1}{2} \\ \frac{3}{4} \end{array}\right\}\ )$

Now we can introduce the lexicon. The entries should be in braces and when there are more than one item they ought to have been arranged vertically, one above the other e.g.,

$\left\{\begin{array}{l} ek \\ bɛ \\ trən \\ etc. \end{array}\right\}$

to indicate obligatory choice of the included elements. We have, instead, omitted the braces, and entered the items in a conventional horizontal line for printing convenience only.

I	:	ek, bɛ, trəṇ, car, p̱āc, chə, sat, aṭh, nəv
X	:	dəs
XI	:	dəs + ek .....nəv
XX(I)	:	dəs.....nevu (+ ek.....nəv)
C	:	sɔ
D^1	:	həjar
D^2	:	lakh
D^3	:	kəroḍ
D^4	:	əbəj

MORPHOPHONEMIC RULES :

pa + ek	⟶	səva	$1\frac{1}{4}$
pa + bɛ	⟶	səva bɛ	$2\frac{1}{4}$
upto			
pa + nəvvaṇũ	⟶	səva nəvvaṇũ	$99\frac{1}{4}$
ərdho + ek	⟶	dɔḍh	$1\frac{1}{2}$
ərdho + bɛ	⟶	əḍhi	$2\frac{1}{2}$
ərdho + trəṇ	⟶	saḍa trəṇ	$3\frac{1}{2}$
ərdho + car	⟶	saḍa car	$4\frac{1}{2}$
upto			
ərdho + nəvvaṇũ	⟶	saḍa nəvvaṇũ	$99\frac{1}{2}$
pɔṇo + ek	⟶	pɔṇo i.e. 'a quarter less'	$\frac{3}{4}$
pɔṇo + bɛ	⟶	pɔṇa bɛ 'a quarter less from two'	$1\frac{3}{4}$
upto			
pɔṇo + nəvvaṇũ	⟶	pɔṇa nəvvaṇũ 'a quarter less from ninety nine'	$98\frac{3}{4}$

Note that of the three fractions pa, ərdho and pɔṇo, the last two, ərdho and pɔṇo are variable adjectives agreeing with gender of the following noun; if the noun is followed by a postposition, they are, as variable adjectives, represented by oblique forms ərdha and pɔṇa These are the rules of Gujarati grammar, applicable to variable adjectives and nouns, and hence they are not discussed here.

T RULE :

(T Rule 1) XX + I	⟶	IXX
e.g. vis + ek	⟶	ek vis
vis + bɛ	⟶	ba vis

etc. with appropriate morphophonemic rules such as :

bɛ + vis	⟶	bavis	(22)
trəṇ + vis	⟶	trevis	(23)
car + vis	⟶	cɔvis	(24)
aṭh + vis	⟶	əṭṭhavis	(28)
bɛ + tris	⟶	bətris	(32)
pāc + tris	⟶	pātris	(35)
bɛ + šitter	⟶	bɔ̃ter	(72)
trəṇ + nevu	⟶	traṇū	(93)

etc. etc. which would cover the numerals from 21–99 (other than the tens, i.e. 10, 20, 30.....90).

Note that the set of numbers 19, 29, 39, 49, 59, 69 and 79 consists of the element ogəṇ : ' one less than '

des + nəv	⟶ ogṇis	' one less than twenty '	(19)
vis + nəv	⟶ ogəṇtris	' one less than thirty '	(29)
tris + nəv	⟶ ogəṇcaḷis	' one less than forty '	(39)
sitter nəv	⟶ ogəṇɛ̃šši	' one less than eighty '	(79)

The set of numbers 89 and 99 consists of the elements 80 + 9 and 90 + 9 respectively.

ɛ̃šši + nəv	⟶ nevvaši	' eighty and nine '	(89)
nevu + nəv	⟶ nəvvaṇū	' ninety and nine '	(99)

We can now take up the base rules (production rules) which produce (i) zero and (ii) synonymous number names.

(i) Integer ⟶ zero

zero : sun, mīḍū

is self explanatory.

(iia) Integer $\longrightarrow \Psi_4 \left(\left\{ \begin{matrix} \Psi_2 \\ \Psi_1 \end{matrix} \right\} \right)$

This rule blocks the ungrammatical number names such as :

* saḍa car sɔ pāc

* səva nəv sɔ sat

When Ψ_4 precedes either Ψ_2 or Ψ_1, the occurrence of Ψ_3 is blocked by the above rule. The above rule produces the strings.

$\Psi_4\ \Psi_2 \longrightarrow$

səva bar sɔ	(1225)
saḍa pɔ̃caši sɔ	(8550)
pɔṇa pəccis sɔ	(2475)

$\Psi_4\ \Psi_1 \longrightarrow$

səva pāc sɔ	(525)
saḍa nəv sɔ	(950)
pɔṇa nəv sɔ	(875)

Note that this rule also blocks ungrammatical numerals

* saḍa tris sɔ

because the absence of Ψ_3 prevents XX(I) to appear here, and Ψ_2 permits only XX + I; thus the tens (i.e. dəs, vis, tris..... nevu) are blocked.

(iib) Integer $\longrightarrow \Psi_2 (\Psi_4) \Psi_3$

This generates strings such as:

pəndər sɔ tris	(1530)
nəvvaṇũ sɔ nəvvaṇũ	(9999)

Note that the absence of initial Ψ_4 blocks ungrammatical strings such as :

*saḍa pəndər sɔ bətris;

but the introduction of (Ψ_4) medially, permits strings such as :

pəndər sɔ saḍa bətris (1532½).

Note that all the strings generated by this rule can also be generated by R_i rule; thus the numbers such as :

pəndər sɔ bətris	'fifteen hundred thirty two'
pəndər sɔ saḍa bətris	'fifteen hundred thirty two and a half'

can also be expressed as :

ek hɛjar pãc sɔ bətris 'One thousand five hundred thirty two'

ək həjar pãc sɔ saḍa bətris 'One thousand five hundred thirty two and a half'

This rule, then, is an optional production rule which allows certain numerals to be stated in an alternative way.

Some comments on the main base rules may be in order.

1. Integer $\longrightarrow$ R_i
2. R_i $\longrightarrow$ $((\Psi_4) \Psi_3 D_i) Ri - 1$

This, as stated earlier, is the recursive rule. R_i can be expanded as R_4 R_3 , R_2 R_1 and R_0 ,R_0 cannot be recursive – it does not appear again on the right. This prevents infinite repetitions of numbers in a string.

3. R_0 $\longrightarrow$ Ψ_1 (Ψ_4) Ψ_3

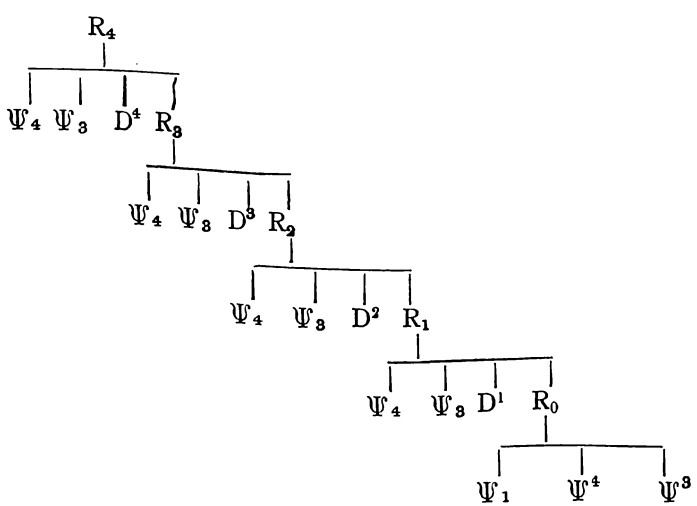
This will block strings such as :

*saḍa chə sɔ pəccis,

by (Ψ_4) placed before Ψ_3 we can get :

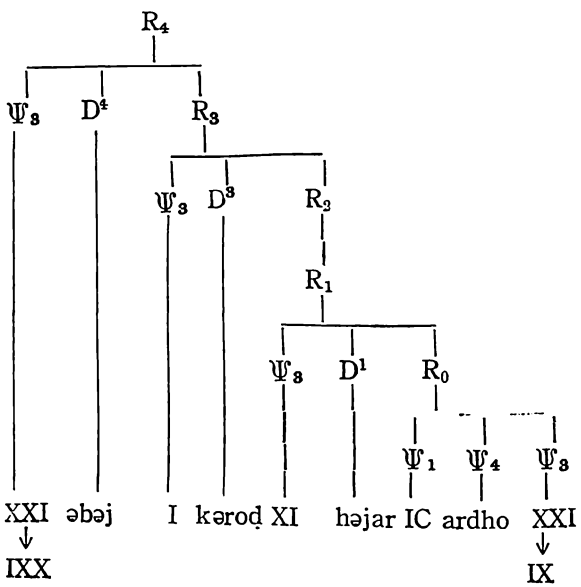
chə sɔ saḍa pəccis ($625\frac{1}{2}$), which is a grammatical numeral.

The above rules of derivation can be illustrated in a tree diagram as follows :



Ψ 1 : hundreds	D^1 : thousands
Ψ 2 : tens + hundreds	D^2 : hundred thousands
Ψ 3 : tens + ones	D^3 : ten million
Ψ 4 : fractions	D^4 : billion

All illustrative examples, i.e. Gujarati number names, would be sub–expansions of this, e.g. ekətris əbəj, trəṇ kəroḍ, ogṇis həjar, saḍa nevyaši (31, 03, 00, 19, 089½).



The step by step application of the above rules to derive the above number name is as follows :

1. Integer $\longrightarrow$ Ri
2. R_1 $\longrightarrow$ $((\Psi_4) \Psi_3 D^1) R_i - 1$
3. R_4 $\longrightarrow$ $((\Psi_4) \Psi_3 D^3) R_3$

   Ψ_3 $\longrightarrow$ XX (1)

T rule 1 : XX (I) $\longrightarrow$ (I) XX

IXX : ekətris

D^4 : əbəj

We have, thus, derived eketris əbəj from $\Psi_3 D^4$. We can now expand R_3.

$R_3 \longrightarrow ((\Psi_4)\ \psi_3\ D^3)\ R_2$

$\Psi_3 \longrightarrow I$

I : trəṇ

D^3 : kəroḍ

We are, thus, working down step by step, from the highest unit to the lower ones. We have now, ekətris əbəj, trəṇ kəroḍ. Expanding the R_2

$R_2 \longrightarrow ((\Psi_4)\ \Psi_3\ D^2)\ R_1$

Items in parenthesis are optional, we take the advantage of this choice and do not expand Ψ_3 or D^2 to the terminal symbols because we do not need them (we do not need any D^2 : lakh, in this number name). We, thus, proceed further.

$R_1 \longrightarrow ((\Psi_4)\ \Psi_3\ D^1)\ R_0$

$\Psi_3 \longrightarrow XI$

XI : ogṇis

D^1 : həjar

Thus, ekətris əbəj, trəṇ kəroḍ, ogṇis həjar.

$R_0 \longrightarrow \Psi_1\ (\Psi_4)\ \Psi_3$

$\Psi_1 \longrightarrow (IC)$

We do not resolve this parenthesis, because we do not require to expand a C.

Note that in all the previous rewrite formulae we chose not to resolve the parenthesis containing (Ψ_4) the fraction. We did not require it. Now, we require it, so we resolve the (Ψ_4) in this rewrite rule.

$\Psi_4 \longrightarrow$ ərdho

Later on, by an application of a morpho phonemic rule we arrive at saḍa in place of ərdho.

$\Psi_3 \longrightarrow (XXI)$

T rule XXI $\longrightarrow$ IXX

IXX $\longrightarrow$ nevyaši.

TAMIL

Non-Terminal vocabulary :

R_0	R_1	R_2	R_3
	D^1	D^2	D^3
	Ψ_1	Ψ_2	Ψ_3
	I	X	X (1)
	C		
	Q		

Terminal vocabulary :

I : oṇṇu, reṇḍu, muuṇu, naalu, anju, aaru, eeḷu, eṭṭu, ombadu (1–9)

X : pattu, irubadu, muppadu, naappadu, aymbadu, arubadu, eḷubadu, emḥadu, tonnuuru (10–90)

X (1) : padinoṇṇu (11)
paṇṇireṇḍu (12)
padimuuṇu (13)
padinaalu (14)
irubattoṇṇu (21)
irubattereṇḍu (22)
irubattumuuṇu (23)
similarly, upto (99)

C : nuuru
D^1 : aayirō
D^2 : laccō
D^3 : koodi

Production rules :

1. Integer $\longrightarrow$ { zero / Ψ 3 / R }
2. Ri $\longrightarrow$ (Ψ_2 Di Q) Ri-1
   $R_3 \longrightarrow$ (Ψ_2 D^3 Q) R^2
   $R_2 \longrightarrow$ (Ψ_2 D^2 Q) R_1
   $R_1 \longrightarrow$ (Ψ_2 D^1 Q) R_0
3. $R_0 \longrightarrow$ Ψ_1 Q Ψ_2 (Q Ψ_3)
4. $\Psi_1 \longrightarrow$ (IC)
5. $\Psi_2 \longrightarrow$ ({ I / X / X(I) })

6. $\Psi_3 \longrightarrow \left\{ \begin{matrix} 1/4 \\ 1/2 \\ 3/4 \end{matrix} \right\}$

7. $Q \longrightarrow \left\{ \begin{matrix} t\,u \\ ttu \\ ee \end{matrix} \right\}$

MORPHOPHONEMIC RULES :

Q functions as a connector. The Q is realised as tu, ttu, or ee according to the differing environments :

nuūru + Q ⟶ nuttu
aayirō + Q ⟶ aayirattu
loccō + Q ⟶ laccattu

Forms such as aayirattu ' thousand..... ' or laccattu ' lakh..... occur only when they are followed by lower numbers e. g.

nuuru + pattu ⟶ nuutt (u) pattu ' one hundred and ten '

aayirō + nuuru ⟶ aayiratt (u) nuuru ' one thousand and one hundred '

laccō + irubadaayirō ⟶ laccatt (u) irubadaayirō ' one lakh and twenty thousand '

Further, the elision of (u) in the above will be stated by the general sandhi rules of the language.

Note that the connector tu is selected only when it connects C, D^1, or D^2 numeral (100, 1000 or 100,000) with some other lower numerals.

We can state lower numerals.

1. $\left\{ \begin{matrix} C \\ D^1 \\ D^2 \end{matrix} \right\} + Q + R_0 \longrightarrow \left\{ \begin{matrix} C \\ D^1 \\ D^2 \end{matrix} \right\} + tu + R_0$

The other connector ee is selected when the numeral koo̱i is used. The following rule states the use of ee.

2. $D^3 + Q + R_2 \longrightarrow D^3 + ee\ R_2$

e. g. pattu koo̱i anji laccō ⟶ pattu koo̱iyee anji laccō : ' ten crores and five lakhs '

oru koo̱di irubadu laccō ⟶ oru koo̱diyee irubadu laccō : ' one crore and twenty lakhs '

ee is also selected when the fractions kaal ' $\frac{1}{4}$ ' and mukkal ' $\frac{3}{4}$ are added :

3. $\Psi_2 + Q + \left\{\begin{matrix}\text{kaal}\\ \text{mukkaal}\end{matrix}\right\} \longrightarrow \psi_2 + \text{ee} + \left\{\begin{matrix}\text{kal}\\ \text{mukkaal}\end{matrix}\right\}$

e. g. anji kaal ⟶ anjee kaal ' $5\frac{1}{4}$ '

pattu mukkal ⟶ pattee mukkaal ' $10\frac{3}{4}$ '

but no connector is required with, are : ' $\frac{1}{2}$ '

e. g. ettu + are ⟶ ettare ' $8\frac{1}{2}$ '

This deletion of Q can be stated as a rule

4. $Q \longrightarrow \phi / -\frac{1}{2}$

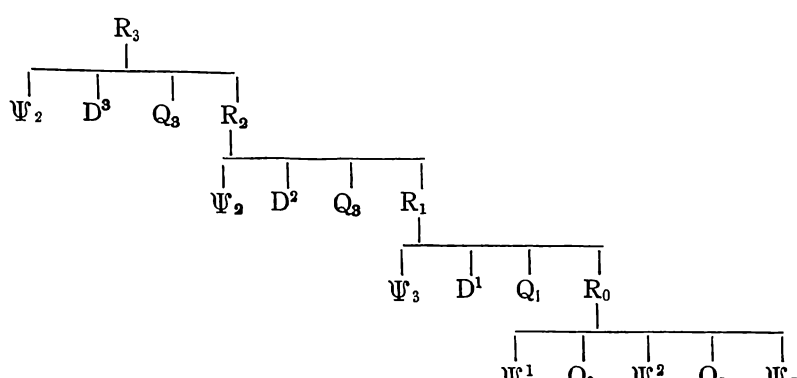
The numbers in the set of tens, 20–90, which have a final –du, replace this –du by the connector –ttu when the unit numbers such as o̱nnu, re̱ndu (1–9) are added. The final –u (of –ttu) is further elided by the general sandhi rules. Thus,

5. X + Q + I ⟶ X ttu I

nappadu + Q + aaru ⟶ nappattaru (46)

aimbadu + Q + re̱ndu ⟶ aimbatterendu (52)

Another suggestion regarding Q, is that morphophonemic rule regarding Q may be eliminated if Q is given subscripts just as D and ψ are. The maximum expansion is :



In this case, the lexical features would be :

ee = Q_3
tu = Q_2, Q_1
ttu = Q_0

The revised rules would be,

2. $R_1 \longrightarrow [\Psi_2\ D^1\ Q_1]\ R_1 - 1$
3. $R_0 \longrightarrow \Psi_1\ Q_0\ \Psi_2\ [Q_3\ \Psi_3]$

The Tamil rules differ from the Gujarati rules in the following way :

i) Absence of such sequences as 'eleven hundred, fifteen hundred' etc. in Tamil obviates the need of a $\psi_2\ \psi_3$ in the base rules of Tamil.

ii) The fractions always come at the end in Tamil numerals. Moreover they are always additive = unlike Gujarati where $\frac{3}{4}$ in a $\Psi_4\ \Psi_3$ i.e. poṇa car is interpreted as 'one quarter less from four : $3\frac{3}{4}$. The complicated rules regarding fractions are not required for Tamil.

iii) Tamil numbers such as twenty one, thirty five etc. (11–99) can be generated from the base rules without any T rule – unlike Gujarati where a T rule XXI ⟶ IXX is required for ekvis (lit. 'one–twenty' etc. (see page 29–30); Tamil numerals (21 – 99) are 'twenty and one' type.

iv) Tamil requires a connector Q in the base rules and morphophonemic rules for the selection of the connector. Gujarati does not require a connector.

We have not formulated two rules for Tamil regarding the set of numbers 10, 20 ..... 90, which can be interpreted as 10, 2 X 10, 3 X 10 ... 9 X 10; the forms pattu : 10; irupadu : (20); etc. can be interpreted as 2 X 10 etc. Another morphophonemic rule required for Tamil is to state the formation of tolayram 'nine hundred' which can be interpreted as 'one hundred less than one thousand; ayaram is 'one thousand' in Tamil.

The Saurashtri rules number names, given below indicate that Saurashtri shares with Tamil all the above four points of difference between Tamil and Gujarati. Saurashtri

and Tamil have, therefore, identical grammar of number names, though the lexicon of Saurashtri is mainly IndoAryan.

SAURASHTRI

Non-terminal vocabulary:

R_0	R_1	R_2	
C	D^1	D^2	
Ψ_1	Ψ_2	Ψ_3	
I	X	XI	XX (I)
Q			

Terminal vocabulary:

I : oṇṭe, diye, thinye, caaɪye, pããcye, sovye, saatye; aaṭhye, novye (1 – 9)

X : desye,

XI : vigaarye, baarye, terye, caudye, phandarye, sovalye, satarye, tiviniviis, oviniviis (also egun viis) (11–19),

XX(I): viis, viis un oṇṭe, viis un diye, ... viis un novye (20, 21, 22 ... ... 29) triis, triis un oṇṭe ... (30, 31 ...), Similarly, caalisye (40), pannaas (50), saaṭ (60), sattar (70), assi (10), desunnɔ oṇṭe sɔ (90). (20 – 99).

C : sɔ (100)

D^1 : sasar (1000)

D^2 : leku (100,000)

Production rules:

1. Integer $\longrightarrow$ { zero, $\Psi_1\ \Psi_3$, Ψ_3, Ri }

2. $R_i \longrightarrow (\Psi_3\ D^i\ Q)\ R_i - 1$

   $R_2 \longrightarrow (\Psi_3\ D^2\ Q)\ R_1$

   $R_1 \longrightarrow (\Psi_3\ D^1\ Q)\ R_0$

   $R_0 \longrightarrow \Psi_1\ Q\ \Psi_2\ (Q\ \Psi_3)$

3. $\Psi_1 \longrightarrow$ (IC)

4. $\Psi_2 \longrightarrow$ ({I, X, XI, XX (I)})

5. $\Psi_3 \longrightarrow$ {1/4, 1/2, 3/4}

Morphophonemic rules:

i) XX+Q+1 ⟶ XX un 1
e.g. viis+onṭe ⟶ viis un onṭe '21'
triis+onṭe ⟶ triis un onṭe '31'

ii) {C, D^1, D^2} +Q+R_0 ⟶ {C, D^1, D^2} +holle+R_0

e.g. onṭe sɔ onṭe ⟶ onṭe sɔ holle onṭe '101'
onṭe sɔ kaal ⟶ onṭe sɔ holle kaal '$100\frac{1}{4}$'
onṭe sasar onṭe sɔ ⟶ onṭe sasar holle onṭe sɔ '1100'
onṭe sasar wonnukkaal sɔ ⟶ onṭe sasar holle wonnukkal sɔ '$1100\frac{1}{4}$'

The above selection rules for the connector are comparable to similar rules for Tamil, though the lexical items (connectors) and the restrictions are different. There is one more differenence: if the first number is thousand, the following connector could be either holle *or* un e.g. onṭe sasar onṭe ⟶ onṭe sasar un onṭe *or* onṭe sasar holle onṭe '1001'; Tamil does not permit such variation.

The highest digit for Saurashtri is sasar 'thousand'; while the highest for Tamil is kooṭi. Note that the digits above thousand in Tamil, leccō and koodi, are loanwords from Indo Aryan.

The four characteristics in which Tamil differed from Gujarati are shared by Saurashtri with Tamil. Thus, in all the following respects, Saurashtri agrees with Tamil:

(i) Saurashtri does not permit sequences such as 'eleven hundred' (as alternating with 'one thousand and one hundred');

(ii) its fractions come at the end and they are additive;

(iii) the numbers 21 – 99 are of the 'twenty and one' type and

(iv) it requires a connector Q.

Saurashtri lexical items for the fractions $\frac{1}{4}$ and $\frac{3}{4}$ are loanwords from Tamil; while the item for $\frac{1}{2}$ is Indo Aryan.

$\frac{1}{4}$: kaal
$\frac{1}{2}$: həddu
$\frac{3}{4}$: mukkal

There are some more morphophonemic rules for Saurashtri fractions which may be stated here.

oṇṭe həddu ⟶ deḍh ($1\frac{1}{2}$)
(cf. Gujarati dɔḍh p. 29)

The numerals 125, 150 and 175 in Saurashtri can be stated by the alternating forms, e.g.

oṇṭi sɔ holle vis un pāāc *or* wonnukkal sɔ
oṇṭi sɔ holle pannaas *or* deḍh sɔ
oṇṭi sɔ holle sattar un pāāc *or* wonnumukkal sɔ;

but unlike Gujarati, these alternating forms are permissible only when the highest digit is hundred e.g. upto 900. 'one thousand five hundred' cannot be translated in Saurashtri by 'one and half thousand' as one does in Gujarati 'dɔḍh həjar'; the permissible Saurashtri form is only 'oṇṭi sasar holle paan sɔ'

The most peculiar feature of the Saurashtri numerals, which is neither found in Gujarati nor Tamil (nor, indeed, in any number system in Indian languages, is the count from 90 to 99. The Saurashtri numeral for 90 is

des unno oṇṭe sɔ 'ten less from one hundred'

Then the count proceeds on the same lines up to 99 in the same fashion.

91 : nov unnɔ oṇṭe sɔ : nine less from one hundred'
92 : aaitye unnɔ oṇṭe sɔ : eight less from one hundred'
93 : saatye unnɔ oṇṭe sɔ : seven less from one hundred'
94 : sɔye unnɔ oṇṭe sɔ : six less from one hnndred'

95 : pããcye unnɔ oṇṭe sɔ	: five	less from one hundred '
96 : carye unnɔ oṇṭe sɔ	: four	less from one hundred '
97 : thinye unnɔ oṇṭe sɔ	: three	less from one hundred '
98 : diye unnɔ oṇṭe sɔ	: two	less from one hundred '
99 : oṇṭe unnɔ oṇṭe sɔ	: one	less from one hundred '

A wholistic approach to language may not explain the Saurashtri - Tamil type of bilingualism. The Saurashtri - historically an IndoAryan language - may have its syntactic structure closely akin to Tamil: large number of utterances of both the language can be generated by the same set of meta-rules.

What are the implications of such a situation ? The grammatical structure of Saurashtri has converged to the Tamil structure; the similarities are noticed at both the levels, phonological as well as grammatical. The process of borrowing is not limited to lexical items or derivative suffixes but it extends to the adaptation of the rules of syntax of the dominant language. The extent of borrowing in Saurashtri has reached a point where Saurashtri - Tamil code-switching for a Saurashtri speaker is performed smoothly in a way comparable to that of the style switching in diglossia; there is, of course, a difference in as much as the code - switching is limited to the Saurashtri community while in diglossia the style switching is common to all the effective communicators in the community.

Tamil is the dominant language in this situation and the direction of borrowing is from Tamil to Saurashtri. Much of the data presented here may be in a formal or literary style of Tamil because the educated Saurashtri informant, under the circumstances, may have liked to display his ability to speak ' correct ' Tamil; the colloquial Tamil in this context may be a different variety. What is significant is the fact that even the formal Tamil and the colloquial Saurashtri of the informant show a considerable degree of structural similarity; in a given context of situation, therefore, the colloquial Tamil and colloquial Saurashtri would be far more closer.

An interesting question may also be posed regarding a review of the notions of pidgin and creole languages; the genesis of mixed languages, a half-way meeting of the barbarian native and the civilized colonial giving rise to a mixed and a stunted and dwarfed language, does not offer a satisfactory explanation and account of the sociolinguistic situation. The idea that a creole ' fills out a pidgin '

also needs reexamination because if we have to use a label for Saurashtri, it would be labelled as a creole, – a creole without a pidgn. Pidgin and creole are broad labels of a contact situation rather than labels of language types or types of language mixtures. Probably many languages pass through such stages and creolisation may be a phase in the life cycle of a language. One of the features of multilingualism or bilingualism in India is its stability; despite the high rate of illiteracy and lack of any tradition of formal language teaching (of Indian languages to non-native speakers) in India, the incidence of bilingualism is significant. The Census data regarding bilingualism indicate that speakers of different languages live side by side, in considerable number in rural as well as urban areas, thereby making a sizable population of each language a bilingual group.

Why are the bilingual situations stable? When a speaker of German or Spanish or Polish migrates to America, he gives up his language after second or third generation. Similarly, immigrants in other countries in Europe give up their languages after a few generations and accept the majority language. The Indian language speaker whether it is Kannada or Panjabi, maintains his speech, no matter where (in India) he settles down or how long he has settled down. In order to settle down among other language speakers, an Indian does not have to give up his language. He is welcome despite his different language; speaking a different language does not make him an alien. The underlying acceptability of any Indian in any Indian cultural setting is symptomatic of a cultural identity and homogeneity at a deeper level; it permits retention of identity markers – whether it is language or religion, food habits or dress habits. Continuous language contact in such multilingual communities results in a set of rules shared by diverse languages, at the same time retaining identity markers, mainly at a morphophonemic and lexical levels. The case studies presented here supply a number of examples of this type of sharing and retention.

Diversity of languages and dialects in multilingual countries have been interpreted as barriers in communication; Ferguson and Gumperz (1960) suggested 'In Asia, however, where intergroup communication is severely limited by ritual restrictions, we would expect these differences to be much greater' (p. 10). I have (1963 and 1967), however suggested that variation should be treated as

functional. The criteria of density of communrication (viz. that the amount of verbal interaction is in inverse ratio to the amount of variation) is not applicable in the understanding of social dialects. Speech variation among different social groups could be symptomatic of communication rather than a barrier in communication because not only it conveys the social status of the speaker and the listener but it facilitates communication so that a higher caste speaker can talk with a lower caste speaker without loosing his 'caste'. Freidreich (1961) also noted that even a pariah loves his speech. Gumperz (1969 b) now considers their early formulation as 'too simple to account for the complexities of the Indian caste situation' and asks some questions : 'The question arises : is it possible to use models derived from European cases to explain the Indian case? Or, is the caste structure so different from the class stratification current elsewhere as to invalidate the application of models derived from non Indian society?' (p. 600). One may note, however, that 'the model derived from European cases', viz. Bloomfield's notions of the density of communication and social dialects may not be applicable anymore to the European cases; it was too mechanistic. Labov's work has shown that we can work with more refined and sensitive models to study speech variation.

Again and again, socio-linguists seem to resort to 'lack of communication' or 'isolation', to explain language diversity. Gumperz, (1969 b) though he informs us that their earlier assumptions were too simple, repeats his earlier notion in the beginning of the same paper (1969 b, p. 597) where he says 'Urbanization is acting as a solvent upon traditionally compartmentalized and localized modes of life'; recent work in Indian sociology has questioned the notion that traditional Indian society was 'compartmentalised and localised'. Recent fieldwork in speech variation has shown that caste-class barriers show sharper demarcation in urban areas rather than in rural areas (Pandit 1969 a).

In his study on Kupwar, Gumperz raises the issue: 'Bilingualism in Kupwar is therefore a long standing tradition. Why is it that it has maintained itself for so long in spite of regular and frequent interaction among local residents? Although it was not possible to collect detailed data on interaction patterns, information obtained from living in the village over a period of several months suggests that the major factor in language maintenance are the local norms

or values requiring strict separation between public and private (intra-kin group) spheres of activity. ' This explanation for a stable bilingual situation is very much similar to the earlier one which was ' too simple ' viz. because there is less communication between groups of people, (those who form a kin-group and the rest) the language of the kin-group is retained. No doubt, neighbourhood and family groups are major factors in maintenance of speech patterns, but men as well as women and children all are active collaborators in a village community; the family or the caste group is not a separate island of communication; this is even more true in urban or semi-urban settings where incidence of multilingualism is even greater. It seems to me that the basic question ' why people maintain their language ' can be answered only in the context of the general acceptability of other language speaking Indians in an Indian cultural setting. An Indian fieldworker, actually, would like to rephrase the question as ' why do people *give up* languages ? '

Grassroot multilingualism and continuous creolisation of speech at colloquial levels pose problems for language teaching and language standardisation. The standard literary varieties superimposed on the local colloquial forms are frequently far removed from the local usage. The pedagogue and the planner should accommodate many intermediate levels of varying norms of selection before textbooks in the standard variety are introduced to the candidates. Considering the broadening base and the democratisation of primary education where speakers from many tribal and backward areas are entering the schools and universities, sudden imposition of a standard literary variety is likely to create a communication gap.

The process of standardisation also deserves watching. On the one hand, speakers transact their daily business in colloquial varieties, which in multilingual areas are capable of communicating across language boundaries, and, on the other hand, the newly formed ' linguistic provinces ' with the powerful language lobbies in the central and state governments have created governmental agencies for instant standardisation of the languages of the eighth schedule. Rapid coinage of ' technical terms ' and rapid production of ' literature '- mainly translations, have further removed the standard varieties from the colloquial usage, and they have turned

more and more to Sanskritic and neosanskritic vocabulary. Interestingly enough, the morphological structure of most of the new standard coinages is that of English and the lexical content – the roots – are Sanskrit (since most of these are morpheme by morpheme translation of English terms).

Multilingual countries have long traditions of communication across languages; problems of communication in multilingual countries are not solved by wishing away the multilingualism or by imposing a dominant language. We should look for the existing channels of communication and put them to good use. An acceptance of these semi-standard local varieties in the sphere of education and standardisation could be a step in that direction.

3. LINGUISTIC EXPRESSION OF SOCIAL DISTANCE

Sanskrit loanwords in some of the New Indo-Aryan (NIA) languages such as Marathi and Gujarati present a type of distribution of consonants which is not distinctive in Sanskrit or NIA taken separately. Intervocalic heterorganic consonant clusters of Sanskrit loan words contrast with intervocalic heterorganic clusters of Gujarati or Marathi: the structure of sanskritic clusters is $c_1 c_1 c_2$ while the structure of the native clusters, Gujarati or Marathi, is $c_1 c_2$; c = any consonant, the underfixed Arabic numerals indicate heterorg anic and homorganic varieties. Following examples show the contrast between the sanskritic clusters of loanwords and the native clusters in the vocabulary of Gujarati and Marathi:*

	Sanskritic clusters	Native clusters
Gujarati:	bhəkkto 'devotees'	bəkto 'chattering'
	pattro 'characters'	patra 'fried fritters'
	vyappti 'application'	vyapti 'spreading'
	uddbhəvyo 'arose'	bədbo 'foul smell'

* Data from NIA languages is in phonemic transcription; Sanskrit and Prakrit words are transcribed according to the known conventions. Item from Nepali Dictionary are transcribed according to the original. Segmental phonemes of Gujarati : ieɛuoɔa pbtdʈɖcjkgsʃxh vyrlḷmnɳN; of Marathi : ieɛuoɔəaiɔuə pbtdʈɖcjkgsʃʂhv yrlḷmnɳŋ ~.
Marathi examples are checked with native speakers. While there is no hesitation among the Gujarati or Marathi speakers in separating the sanskritic clusters and native clusters whose second element is y or v, I have noticed a difference of opinion among Marathi speakers whether they distinguish native clusters with two stop consonants from sanskritic clusters. They agree that 'sometimes' they perceive the difference; moreover, they do not agree that the difference is related to the difference in context. Dr. Ashok Kelkar informs me that 'in Marathi the c_1y– c_1c_1y contrast is valid for all social levels and for all speech styles. What you say is true, however, of other contrasts : i-i., i-i̯, s–š–ʂ; c_1c_2– $c_1əc_2$ etc.' Such observations could be appreciated if it is borne in mind that the difference is highlighted only when the context is learned and very formal. As a native speaker of Gujarati, and as one belonging to the 'learned' group, I am quite aware of the difference, and I perceive the difference in the speech of others also.

	bhəvvyo ' pious souls '	avyo ' came '
	appto ' venerable elders '	apto ' giving '
Marathi:	puṇṇyala ' for meritorious ' deed '	puṇyala ' to Poona '
	suppte ' asleep '	khuptə ' pricking '
	dhattvadeś ' ordering of verb-roots '	satva ' seventh '
	sattvik ' wholesome '	natvala 'to the grandson'

A sanskritic cluster can be described as a cluster whose elements are borrowed from Sanskrit; if one of the elements is a NIA evolute, then the cluster is not a sanskritic cluster, it is a native cluster. Thus, in a Sanskrit loanword bhəkkto, the sanskritic source is Skt. bhakta – ' devotee ', so the cluster kkt in the loanword is sanskritic, the final o is Gujarati plural suffix. In uddbhəvyo' the sanskrltic source is Skt. udbhav – ' to emerge ', so the cluster ddb in the loanword is sanskritic, but the cluster vy in the same word is not sanskritic, here, the element y is non-sanskritic, it is the past tense suffix of Gujarati; o is the masculine nominative suffix of Gujarati; therefore, vy is a native cluster. In bhəvvyo, the sanskritic source is Skt. bhavya– ' grand ', so the cluster vvy in the loanword is sanskritic, the final o is the plural suffix of Gujarati. The native clusters on the right column, apto, vyapti, puṇyala etc. do not have a sanskritic source consisting of pt, ṇy etc. ; the native clusters are, of course, noticed in borrowings from languages other than Sanskrit, such as English or Persian. The non-initial sanskritic clusters are distinctly marked by a closure of longer duration.

These sanskritic clusters appear in the speech of the educated elite and are more prominent in formal contexts. In informal contexts, the sanskritic clusters are apt to be realised as native clusters. The uneducated hardly employ the elite sanskritic vocabulary, and if they have an occasion to employ the sanskritic clusters, they usually realise them as native clusters. Similarly, the educated elite belonging to non-sanskritic tradition such as the Zoroastrian or Islamic, realise the sanskritic clusters (they do employ the sanskritic vocabulary in formal contexts) as native clusters.

The structure of sanskritic clusters in native vocabularies attracts attention because the elements of sanskritic clusters are

noticed neither in the Sanskrit phonemics as we know it from traditional sources nor in the phonemics of NIA languages. We also do not know whether the sanskritic clusters in native vocabularies have always displayed this distinction or whether they once conformed to the Sanskrit phonemic pattern and the change took place later. The mechanism of such change may be borrowing with status groups in the community acting as catalysts of change. We have attempted here to enquire into these related problems to explain the distinctive structure of sanskritic clusters.

We started with the assumption that $c_1 c_1 c_2$ types of consonant clusters are not distinctive from $c_1 c_2$ types in Sanskrit (later on, we will enquire into the validity of this assumption); since the Sanskrit loanwords appear in NIA languages with $c_1 c_1 c_2$, our first question is : at what stages of the history of Indo-Aryan (IA) languages do these loanwords appear? It is fairly easy to separate a Sanskrit loanword borrowed in post- Middle Indo-Aryan (MIA) period, because one of the significant features of MIA phonology is the absence of heterorganic consonant clusters; (consonant clusters with r as the second element have beea retained in Western and North-Western dialects of MIA); the clusters of old Indo-Aryan a (OIA) are assimilated in MIA, regressively if the second element is a stop and progressively if the second element is a spirant, with palatalisation and labialisation if y and v respectively are second elements. If a NIA word, appearing with a heterorganic consonant cluster, corresponds to a Sanskrit word with a heterorganic consonant cluster then it is a Sanskrit loanword borrowed in post-MIA period. But if the loanword belongs to early MIA period, we can detect it only by basing our inferences on the phonemic systems of MIA at different stages.

The MIA period represents the split of Indo- Aryan speech community into various dialect groups. The split started very early, the dated evidences of regional variations go back to the inscriptions of emperor Ashoka, around third century B. C. In this early phase of MIA period, the literature is divided in two branches, the Jain and the Buddhist. Later on, MIA develops as a style, and survives up to 12 th or even 14th century A. D. During this period, Sanskrit (representing OIA) also survives as a style. Though both have a parallel existence, there is a noteworthy difference: phonemic and grammatical structure of post-Pāṇinian Sanskrit is maintained

unchanged in literary Sanskrit throughout its parallel existence with the developirg Indian languages; MIA, on the other hand, does not present a uniform system, its links with the changing Indian languages are made obvious through the changing phonemic and grammatical patterns of MIA. Sanskrit is held up as a norm of purity, immutable, While no such claims are made for the Prakrits.

One of the significant changes in the phonemic patterns of MIA of different periods is relevant to our inquiry. The distribution of stop consonants in early MIA and late MIA is different. Though the areal and temporal gradient cannot be very clearly demarcated, one can state with available inscriptional and literary evidences that in early MIA, consonant gemination is contrastive, single and geminate consonants occur medially, while, in late MIA they are in complementary distribution. In the NIA cognates for 'lotus' for example : Gujarati poyṇi, Sinhalese piyuma, Kumaoni payā and Nepali paīyū (R. L. Turner, 1931 s. v. paīyū) one can see the loss of consonant of the original cluster in Skt. padma–. Sindhi pabiṇī, however, indicates that there is no intermediate stage of simplifying a consonant cluster by an epenthetic vowel, it is the result of the common MIA sound change, an assimilation of heterorganic consonant cluster. These two types of cognates suggest that while the Sindhi item is in direct line, through an intermediary sound change, of Skt. padma –, the rest of the cognates are a development of a loanword from Skt. padma –, borrowed in early MIA as paduma –. Many of such early loanwards in MIA, with epenthetic vowel in the original consonant cluster, are recorded in Prakrit; (Pischel, 1900; § 132) others have to be inferred.

Another type of loanwords which can be inferred from the MIA vocabulary are words with geminated consonants whose reflexers in Sanskrit have single consonants. NIA cognates for 'one' for example : Gujarati, Hindi and Marathi ek, Bengali ɛk, with the retention of original single intervocalic k of Skt. eka–, indicate that it is a loanword from Sanskrit. Assamese alternants e ~ ek and Nepali yeu-ṭā indicate that e and yeu– are direct developments of Skt. eka– with the loss of intervocalic stops in the late MIA period (Turner 1931 s.v. yeuṭā; Kakati 1941 § 386a). Sanskrit words borrowed during the period when single stops did not occur intervocalically are substituted by geminated stops and conform to the MIA pattern. Numerous words of this type are noticed in

Prakrit literature. The alleged influence (Pischel 1900 § 194) of accent on gemination e.g. Skt. niyamá– > Pkt. nimma 'regulation', Skt. jitá – > Pkt. jitta 'conquered', Skt. tailá– > Pkt. tella 'oil', cannot be sustained because there are many more words where the single intervocalic stop is lost inspite of the accented syllable e.g. Skt. pītá– > Pkt. pīa 'drunk', Skt. kṛtá– > Pkt. kaya 'accomplished' etc. etc. Moreover, our knowledge of the accentual system of MIA is so meagre that it can hardly be invoked to support any argument.

The two waves of loanwords from Sanskrit in the early and late MIA periods examined above give us only a negative answer to our main problem; Sanskrit words with heterorganic consonant clusters could not have formed part of MIA vocabulary. If they were borrowed in early MIA period they would have been simplified with an epenthetic vowel and ultimately the consonantal element would be lost; if they were borrowed in the late MIA period the clusters would be assimilated to a homoraganic geminate and eventually survive as single stops (or as geminates in the languages of the North-Western group, such as Sindhi-Kacchi and Panjabi-Lahanda). Our classification of these loan types is restricted to the loanwords from one literary medium to another literary medium; loanwords from Sanskrit, when they appear in Prakrit, always conform to the Prakrit phonological pattern. Prakrit writers were writing in literary media and their native dialects or Sanskrit were distinctly kept separate. This resulted in a very conscious Prakritization of the loan elements in Prakrit literature. Prakrit grammatical tradition (fairly well established during 500–1000 A.D) is mainly concerned with rules of converting Sanskrit to Prakrit, and any Prakrit writer, to be sure, must have undergone considerable schooling in Prakrit, before he started composing. The actual situation, during the MIA period, must be quite different. The Sanskrit knowing elite must have always employed a number of Sanskrit words, even though the speakers belonged to some MIA speech community. The prestige attached to Sanskrit words would have prevented them from conforming fully to MIA speech patterns and such words may have continually appeared as elegant variants of the innovating elite dialects. The heterorganic consonant clusters in Sanskrit words e.g. pt, kt, vy, etc. could have appeared in their Sanskrit shape as

elegant variants or with an epenthetic vowel between the cluster in early MIA and with an assimilated homorganic geminate in late MIA. Up to this stage, we do not find any situation which would permit us to posit the appearance of Sanskrit clusters as ppt, kkt, vvy etc. as loan elements. Selection of pt, kt, vy etc. in one's speech was enough to convey elegance and refinement. There is a me hre aaarky tspethe mark chancet turopn eddtfiph eo vrian cgip ($c_1c_1c_2$) etc. instead of pt(c_1 c_2) etc.; even if such forms may have appeared as variants, their systemic status in Sanskrit as spoken by the elite could only be as non-distinctive variants of pt (c_1 c_2) etc. types; their status in the innovating dialects of MIA could also be interpreted as c_1 c_2 only. Chances for these clusters to appear as distinctivc c_1 c_1 c_2 in the innovating dialects of MIA are meagre, because, clustering of heterorganic consonants is foreign to the speakers of MIA dialects. It may be irrelevant to speculate about the phonetic realisations of c_1 c_2 in the innovating dialects of MIA, but we can take into consideration that there is no divergent development of c_1 c_2 and c_1 c_1 c_2 types in post-MIA period; hence conditioned variants need not be established. Of course, if the elegant variants are interpreted as learned loanwords borrowed anew each time, and rightly so, then the question of divergent development does not arise.

NIA period alone, now, could offer an explanation; impact of the phonemic patterns of NIA, at different stages, could be responsible for the altered nature of sanskritic clusters. The NIA period, from the point of view of the present problem, can be divided in two stages. In early NIA, single and geminate consonants contrast; single consonants occur intervocalically (languages of the North-Western groups Panjabi, Lahanda, Sindhi, and Kacchi may have branched off earlier). This makes it possible for the Sanskrit loanword with a heterorganic cluster to be substituted in NIA with an epenthetic vowel intruding between the cluster. Thus, Sanskrit words such as rakta – 'tinged', gupta– 'secret', svapna– 'dream' etc. are realised in early NIA as rəgət (~ rəkət), gupət, səpən etc. (appearance of single intervocalic stops as distinct phonemic units in early NIA is similar to the situation in early MIA; Sanskrit loanwords in early MIA also appear with epenthetic vowels. In NIA, however, the epenthetic vowel is mainly ə). In the later NIA, by a change in the accentual system, final ə is lost, and still later, penul-

timate ə is lost (for the loss of ə in NIA languages, see : Bloch 1920 § 38; Saksena 1937 § 103; Turner 1931, Pandit 1961 Lang. 37.64). This later change, in a number of NIA languages, brings together two heterorganic consonants separated by ə : Gujarati bəkəto > bəkto, apəto > apto; Hindi upəḍa > upḍa 'lifted', əbhagənī > əbhagnī 'unfortunate woman' etc. (here again, Sanskritic pattern of heterorganic consonants is recreated in the later phase of NIA languages). Furthermore, in languages like Gujarati, vowel sequences of the type iu, ia, ua and uo developed as yu, ya, va and vo and sequences of the type c +y or c+v in avyo 'came' and kəḍva 'bitter' replaced earlier (old Gujarati) āviu and kaḍuā. From this late NIA stage upto the present stage there has been no phonemic barrier to the borrowing of Sanskrit words with heterorganic consonant clusters in the native languages. Yet, the clusters of native inherited vocabulary are realised as c_1 c_2 while the clusters of Sanskrit loanwords are realised as c_1 c_1 c_2 and the contrast is maintained in elite idiolects and formal situations.

These sanskritic words, hitherto borrowed at various stages of the development of Indo-Aryan languages, were marginal to the native systems. The sanskritic pattern was foreign; the loan elements survived in the innovating dialects as learned borrowings; but, when the NIA languages developed in the direction of creating favourable phonemic patterns which would permit the entry of Sanskrit loanwords, the sanskritic pattern in the loanwords of innovating dialects altered and maintained the distance from native patterns. The source of sanskritic clusters cannot be sought in the interlanguage substitution at the phonetic or phonemic level, it is an innovation of the elite, result of a tendency to maintain social distance through linguistic expression.

Status significance of learned words is maintained by constant borrowings and it is possible that the literary tradition and script may largely contribute to determine the distinctive shape of loan element. In Sanskrit, consonant clusters are written as clusters, i. e. the first consonant is half, the second consonant is complete ('complete' indicates syllabicity 'half' indicates the lack of syllabic element,, i, e. the vowel). Optionally it was permissible to write a geminate consonant after r, in the same syllable, i. e. a word like svarga – 'heaven' could also be written as svargga–.

The native clusters in the NIA languages languages are not written as clusters. The graphic traditions were established prior to the operations of sound changes which brought about the existence of clusters in modern languages. Thus, modern Gujarati šəkti or modern Hindi səkti 'being able to' is not written like Sanskrit šakti – 'energy'; the graphic tradition interprets the native words šəkti or səkti as having three syllables šə-ke-ti, sə-ke-ti, and the Sanskrit word having two syllables šak-ti, (with k written as half). The graphic traditions keep the native clusters and the sanskritic clusters apart. Even here, in the cases of clusters with second element as y, sanskritic clusters and native clusters are written alike. Thus, in Marathi, puṇyala 'to Poona' and puṇṇyala 'for meritorious deed' are written alike, with half ṇ and complete y; similarly in Gujarati, pətyo 'finished' and sə̄ttyo 'truths', the clusters are written alike with half t and complete y. The realisation of sanskritic clusters cannot be called a spelling pronunciation, because, the sanskritic clusters are spelt as c_1 c_2 only. One can state that the distinction between native clusters and sanskritic clusters is maintained in spelling as well as in speech; it is possible that the script may have supported the tendency to keep prestige patterns distinct from native patterns but the distinction in speech is achieved only by innovation.

We have assumed that clusters in Sanskrit were of the c_1 c_2 type only; this is based on the evidence of Sanskrit phonemics; contrast of the type c_1 c_1 c_2 with c_1 c_2 is not noticed in Sanskrit. Though we have no way to get at the phonetic data of that period, early Sanskrit treatises on phonetics have described such clusters as those in which the first consonant is in 'close contact' with the following stop (Rk-prātiśākhya vi. 17–8; Allen 71–3; 1951). Phonetically, therefore, the consonant clusters can be interpreted as c_1 c_2 only.

The sanskritic clusters are illustrative of pressures exerted by social distinctions on linguistic change. Sanskrit words in the idiolects of NIA languages are markers of prestige. In many cases, not the whole word but only some elements are sanskritic (for an analogous situation in Nepali, see Turner 1931). Thus, in Gujarati, in a word such as kuḷ 'family' the sanskritic element is u; this element is a learned borrowing from Sanskrit kula- 'family'. Here, while l > ḷ is a regular sound change, the expected sound change u > ə

(*kəḷ would be the expected regular development) is shut out by the learned influences which are ultimately derived from literary or oral Sanskrit traditions.

Standards of elegant speech are connected with literary tradition and spelling pronunciation; they are superimposed on regional dialects. In Gujarati, for example, the standard colloquial is a regional variety, but still the image of a 'pure' standard language is largely formed by spelling pronunciation and sanskritic elements. In standard colloquial Gujarati, for example, there are two phonemes s and š; there is, however, a post-velar [whose range is from post-velar to prepharyngeal] spirant x, noticeable in the standard colloquial; its frequency is proportionate to the formal-informal situation, elegance and literacy of the speaker. x is a low-prestige element, but it occurs in free variation with s and š in standard colloquial; in low-prestige dialects and informal contexts it occurs more frequently; it replaces s in Sanskrit loanwords also. But, in standard colloquial, x does not replace s in sanskritic words which are more literary; e.g. in my speech, s is never replaced by x in a loanword such as Skt. asatya- 'untruth'; in my speech, s and x are not really free variants [s-x variation is not inconsequentlal, it has a status significance], they should be interpreted as two phonemes. Interpreting the situation from the point of view of linguistic change, one can say that the phoneme s of the earlier stage is split, some of the allophones have merged with š and the rest have moved to x [or, tend to move to x]; but, because the script has only two symbols s and š, high prestige idiolects continuously borrow s from literary sources. At the same time, it is noticed that in some of the neighbouring dialects, upper caste and lower caste idiolects cannot be distinguished by presence or absence of s; they depend on some other sanskritic element to signify the status difference.

Ever since Jules Bloch's paper 'Caste et Dialecte en Tamoul' in 1910 and more so after the recent impact of linguistics in 1950, linguists interested in the observation of speech variation have tried to correlate social dialects with the caste-system. A number of articles in 'The Linguistic Diversity in South Asia' deal with such correlations and the varieties are commonly labelled as 'caste dialects'. The field workers started with an assumption that 'the Indian caste system makes for the easy recognition of social levels with which linguistic variation is correlated'. Having established

'caste-dialects', impressionistically, – which was not very difficult, because they started with caste as a relevant linguistic group – some conclusions were drawn regarding channels of communication in 'South East Asian' languages. The editors of the Linguistic Diversity in South Asia observe 'The existence within a speech community of social distinctions such as those of caste, class, professional guild therefore give rise to differential rate of linguistic change' .....(p.9). 'In Asia, however, where inter-group communication is severely limited by ritual restrictions we would expect these differences to be much greater' (p. 10).

This reasoning, probably based on the Bloomfieldian notion of density of communication (where variation is interpreted to be in inverse ratio to the amount of verbal interaction) is weak in as much as it assumes that :

(i) Variation is symptomatic of a barrier in communication.
(ii) Caste is the only relevant speech group.

It would be more valid to interpret variation as functional, i.e. it may facilitate communication, and it may act as carrier of social information.

A linguistic expression of social distance can appear only when there is an interaction between low prestige speaker and a high prestige speaker. Changes such as replacement of c_1 c_1 c_2 for c_1 c_2 and s for x in high prestige idiolects presume a continuous verbal interaction with law prestige idiolects. Display of refined speech is spurred on when communication with these who lack refinement is densest. Such variations are not diagnostic of weak lines of communication. On the contrary, one might speculate that the sanskritic elements borrowed by the ascending castes and spread through a process of imitation to various low-prestige idiolects with varying degrees of 'correctness', provided a useful channel of communication between castes; a high prestige speaker can communicate with a low prestige speaker and still maintatn his status by introducing sanskritic elements in his speech.

From this diachronic evidence of linguistic expression of social distance we now turn to a synchronic study of the parameters of variation.

4. PARAMETERS OF SPEECH VARIATION IN AN INDIAN COMMUNITY*

Systematic descriptions of speech variation are recent; both the comparativists and the descriptivists used to ignore the variations with an elegant denial to touch anything 'extralinguistic'. Variations appear as variations only when one is building a micromodel of language structure–the search for constants–and discarding the variants to be appended or brushed aside. But as soon as the linguist learns to outgrow the units of his micromodels and takes an overview of the macrostructure, he may find that what appeared to be variants are structured within a social system. It has been customary to say that greater the verbal interaction lesser the speech variation; that variation marks off boundaries of communication in a speech community. In a model with homogeneity as the underlying assumption there were no barriers other than the barriers of geography to cut off or to restrict communication. This was true within this model, but it was dangerously wrong to assume away the heterogeneity in society; the social barriers of status have ingenious ways of communication without assimilation. Language is not merely a habit, it is a conscious act of distinctive behaviour. Thus, in a given cultural context, one views variation as functional. After all, variation helps preserve one's status; a brahmin can talk unconsciously with others only in his own group; he consciously displays some social markers to help him retain his special identity. Moreover, speech variation offers an excellent vertical channel of social mobility and recognition. The linguist is not the only beneficiary here; the insights thus obtained are for the area students as well. The status seekers, the new aspirants, highlighted by the linguist, cast doubt on the old versions of stratified equilibrium in a caste system.

---

* The fieldwork for the study was conducted for me by my students Dr. Shantilal Acharya (Gujarat Vidyapeeth), Dr. Yogendra Vyas (Gujarat University), and Miss Anjani Kavi (Gujarat University). The statistical analysis was carried out by my colleague Dr. A. G. Ajgaonkar (Deccan College). I am indebted to all these collaborators.
This study and the fieldwork were conducted without any grant or financial aid.

Linguistic diversity in India has attracted some attention and a number of studies have appeared in the last decade; it was almost inevitable that these linguists took for granted the caste system as a structure which intervenes between language and speech in an Indian community. But the caste is not the only parameter for a sociolinguistic inquiry, nor should caste be interpreted as a closed group. The fieldworker starting with a selection of linguistic variables and plotting the differences among different groups in an urban or a nonurban environment may be able to present a more meaningful picture of social classes than the one who starts with a readymade scale of caste and plots the variants along that scale.

I have tried to present certain significant parameters of speech variation in Gujarati; this study was planned in 1961 and the field work was completed in 1963.

The linguistic variables selected in this inquiry are :

1. Nasalisation or the loss of nasalisation of the word final and utterance final ã and ũ.
2. Replacement of initial d– by ḍ– in words where the following syllable has a retroflex consonant.
3. Interchange and selection of the allophones ϕ f and ph of the phoneme / ph /.

The inquiry was conducted in Ahmedabad (Gujarat capital, pop. 1 1/4 million, urban) and in Surendranagar (70 miles north-west of Ahmedabad, pop. 50,000, town,nonurban); the random sample of the population consisted of literate and illiterate men and women and members of the elite caste – Nagar brahmins. The data are processed by standard statistical techniques and the correlations specified are within confidence limits.

Nasalisation :

Gujarati has an eight vowel system

i		u
e	ə	o
ɛ	a	ɔ

Of these, ĩ, $\tilde{\varepsilon}$, ã, ũ, $\tilde{ɔ}$ can occur as nasalised vowels (there is no contrast between nasalised $\tilde{e}$ – $\tilde{\varepsilon}$ and nasalised $\tilde{o}$ – $\tilde{ɔ}$). There are further restrictions on the occurrence of the nasalised vowels : ĩ, $\tilde{\varepsilon}$ and $\tilde{ɔ}$ do not occur in final position in polysysllabic words :

vīṭi	'ring'	phɛ̃ṭo	'turban'
cī	'chirping'	bhɔ̃	'honking'
bhīt	'wall'	bhɔ̃ko	'piercing'
bɛ̃	'bleeting'		

This leaves out two vowels ā and ū which could occur as nasalised vowels in final as well as nonfinal positions :

šū	'what'	kā	'why'
khūṭo	'(tethering) post'	kāṭo	'thorn'
khaṭū	'sour' n.sg.	khaṭā	'sour' n.pl.

We have investigated the variation in the nasalised and the nonnasalised final ā and ū. The difference about the loss of nasalisation in the word-final and utterance final posititions is made because it was noticed that words with nasalised vowels before juncture i.e. final vowels (all vowels have longer allophones before juncture) alternate with nonnasal vowels if they are not prejunctural. This is true not only of the polysyllabic words but also of the monosyllabic words, thus :

šū	vs	šuche (che is the auxiliary verb)
khaṭā	vs	khaṭachē

Accordingly, in this investigation, the distinction labelled as medial -ā-, final -ā, medial -ū- and final -ū refers to prejunctural and final vowels respectively. It should be noted that the word final ū or ā in polysyllabic words or in words of CVw (w = semivowel w) structure always marks a morpheme boundary and is a morpheme: -ū nom. sg. n. or Ist. pers. sg.
-ā nom. pl. n.

The Gujarati spelling rules require the marking of the nasalised vowels. A historical background of the changes in nasalised vowels in final positions may indicate the direction of change.

OG -ai, OG -aī > MG -e but OG -au > MG -o; OG -aū > MG -ū

Paradigm of the verb kar- 'to do' present. indicative in OG and MG is as follows :

	OG		MG	
	sg.	*pl*	*sg.*	*pl*
1st pers.	karaũ	karaũ	karũ	(karũ ~ karie is an analogical development)
2nd pers.	karai	karau	kare	karo
3rd pers.	karai	karaĩ	kare	kare

This indicates that the final -u and -ũ were phonemically distinct in OG while final -i and -ĩ were not distinct (at least in one environment).

Thus, during the history of the last one thousand years of Gujarati language, certain significant changes in the system of the nasal vowels can be noticed. One of these is a gradual loss of oral-nasal distinction in the vowels in the final position in polysyllabic words. Of these, ĩ / -i nasal -oral contrast ceased to be distinctive in the old Gujarati period (i. e. upto about 15th century A.D.), but -ũ / -u were distinctive; -ã and -a were also distinctive all throughout the OG - MG period. The present inquiry indicates the current trends : retention of the nasal vowel -ã / - ũ is a significant expression of social distance, and the upper castes tend to retain it, while the rest - literates and the illiterates - definitely tend towards non-retention of the final nasal vowels. The various parameters of education, urbanity (and income etc.) may have a significant bearing on the social ranking.

The questionnaire for the inquiry of these two variables among the literate and the illiterate in the urban area consisted of one hundred questions out of which there were eleven items for which a nasalised vowel (ũ or ã) was expected; other questions contained minimal or near minimal pairs; if sacã 'true' n.pl. and sacãche 'true' n.pl. with auxiliary were the words in the sentences, there would be two other sentences with the words saca m.pl. and sacache m.pl. with auxiliary, in the questionnaire. The fieldworker uttered a sentence and asked the informant to repeat it after him in his normal casual style. The fieldworker had just to note on a sheet of paper whether he heard the nasalised vowel in the sentence. The informants did not know the purpose of the inquiry. The questionnaire was administered to fifty illiterates and fifty literates (an illiterate is defined as one who does not recognize any written symbol

and a literate as one who can at least read a newspaper headline). The random sample of 50 literate and 50 illiterate from Ahmedabad consisted of men and women who had a continued residence in Ahmedabad; they came from all castes though it is apparent that among the illiterates there is a preponderance of lower castes.

A second sample (population 25 and frequency of the nasalised item 5) was obtained later, from Ahmedabad and Surendranagar; the three groups investigated were illiterate, literate and Nagar; the Nagars are upper among the urban Brahmins and illiteracy among them is almost nil; informants from all the three groups were women and the fieldworker was also a woman.

A third sample (population 25 and frequency of the nasalised item 10) was obtained from Surendranagar; the three groups investigated were mixed (men and women) illiterate, literate (a small group of 13 only) and Nagar; the word final and utterance final distinction was omitted from this sample and only the word final -ā and -ũ were investigated from this area,

1. -ā- (tables 1-8):

URBAN (MIXED)

[-ā-]

(1) Illiterate		(2) Literate	
Speakers	Frequency	Speakers	Frequency
8	zero	1	0
17	1	9	1
10	2	13	2
7	3	8	3
1	4	6	4
2	5	5	5
1	6	2	6
2	7	3	7
1	8	1	8
1	9	1	9
		0	10
		1	11
			12
50			

Mean = x	— 2.18
Variance = s^2	— 4.5476
ns^2	— 227.38 = 50 x 4.5476
$\frac{ns^2}{n-1}$	— 4.6404

Mean	— 3.38
Variance	— 5.3556
ns^2	— 267.78
$\frac{ns^2}{n-1}$	— 5.4649

Population : 50
Frequency : 11

Population : 50
Frequency : 11

n = sample size = 50

$\frac{n}{n-1}$: population correction factor : pcf.

s^2 = variance

$\sqrt{s^2}$ = $\sqrt{\text{variance}}$ = standard deviation

URBAN (WOMEN)

[-ā-]

(3) Illiterate		(4) Literate		(5) Nagar	
Speakers	Frequency	Speakers	Frequency	Speakers	Frequency
8	zero	12	zero	2	1
7	1	10	1	2	2
6	2	3	2	10	3
3	3	—		6	4
0	4	25		5	5
1	5			—	
—				25	
25					

Mean	— 1.32
Variance	— 1.5776
ns^2	— 39.44
$\frac{ns^2}{n-1}$	— 1.6433

Mean	— 0.64
Variance	— 0.4704
ns^2	— 11.76
$\frac{ns^2}{n-1}$	— 0.49

Mean	— 3.40
Variance	— 1.28
ns^2	— 32.00
$\frac{ns^2}{n-1}$	— 1.3333

Popularion : 25
Frequency : 5

Population : 25
Frequency : 5

Population : 25
Frequency : 5

NON-URBAN (MIXED)

[-ã-]

(6) Illiterate		(7) Literate		(8) Nagar	
Speakers	Frequency	Speakers	Frequency	Speakers	Frequency
23	zero	11	zero	3	zero
1	3	1	1	1	2
1	7	1	2	1	3
—		—		4	4
25		13		2	6
				3	7
				4	8
				3	9
				3	10
				—	
				24	

Mean	— 0.40	Mean	— 0.2308	Mean	— 5.9583
Variance	— 2.16	Variance	— 0.3313	Variance	— 10.1237
ns^2	— 54.0	ns^2	— 4.3075	ns^2	—242.9688
$\frac{ns^2}{n-1}$	— 2.25	$\frac{ns^2}{n-1}$	— 0.3590	$\frac{ns^2}{n-1}$	— 10.5639
Population : 25		Population : 13		Population : 24	
Frequency : 10		Frequency : 10		Frequency : 10	

(a) The literate-illiterate difference is not significant in the non-urban area; both the groups drop the nasalisation.

(b) The literate-illiterate difference comes at 5 percent confidence level in the urban area.

(c) -ã- illit. shows a low mean and a low variance, with most of the speakars bunching in the low frequency range while in -ã- lit. there is a greater scatter, and middle frequencies are also filled in.

(d) In one urban sample of women informants, however, though illit. and lit. women are largely bunched in the low frequency range, there is a greater scatter in the illit. -ã-; the mean and variance in illit. women are also higher than those in lit. women; but this difference, again, does not come within confidence level.

(e) The Nagars in the non-urban and the urban area display a distinctly different behaviour; though there is some difference within the Nagars – the nonurban and the urban. The urban Nagars show a high mean and low variance, with more speakers in high frequency range, the non-urban Nagars show a high mean and a high variance. The Nagars, urban or non-urban, definitely show a trend at the retention of -ā- and the rest, lllit, lit., urban-non-urban definitely show non-retention of -ā-.

Utterance final -ā : (tables 9 – 13) :

URBAN (MIXED)

[-ā]

(9) Illiterate		(10) Literate	
Speakers	Frequency	Speakers	Frequency
3	zero	2	zero
4	1	4	1
8	2	4	2
5	3	11	3
13	4	5	4
3	5	6	5
4	6	1	6
3	7	6	7
4	8	5	8
1	9	2	9
1	10	2	10
1	11	2	11
50		50	

Mean	—	4.18
Variance	—	6.7876
ns^2	—	339.38
$\frac{ns^2}{n-1}$	—	6.9261

Population : 50
Frequency : 11

Mean	—	4.86
Variance	—	8.4404
ns^2	—	422.02
$\frac{ns^2}{n-1}$	—	8.6127

Population : 50
Frequency : 11

URBAN (WOMEN)

[-ā]

(11) Illiterate		(12) Literate		(13) Nagar	
Speakers	Frequency	Speakers	Frequency	Speakers	Frequency
9	zero	14	zero	1	zero
4	1	6	1	3	1
4	2	3	2	4	2
8	3	0	3	5	3
2	4	1	4	5	4
2	5	1	5	7	5
29		25		25	

Mean	— 1.8621	Mean	— 0·84	Mean	— 3.24
Variance	— 2.5326	Variance	— 1.6544	Variance	— 2.2624
ns^2	—73.4448	ns^2	— 41.36	ns^2	— 56.56
$\frac{ns^2}{n-1}$	— 2.6230	$\frac{ns^2}{n-1}$	— 1.7233	$\frac{ns^2}{n-1}$	— 2.3567

Population : 25 Frequency : 5

Population : 25 Frequency : 5

Population : 25 Frequency : 5

Only the urban area data are available.

(f) -ā illit. shows a trend towards > a; more speakers are in the lower frequency range; there is low mean and low variance.

(g) -ā lit. shows the same low mean but there is somewhat greater variance, with a somewhat greater scatter. Though both the lit. - illit. show trend towards non-retention, the difference in the behaviour of lit. - illit, with reference to -ā is at 5 per cent confidence level.

(h) Like -ā-, here also a sample from women informants show low mean and variance among the lit. and relatively higher mean and variance among the illit; but the difference does not come within any significance range.

(i) Nagars show high mean and low variance; more speakers are in the higher frequency range; definitely show retention.

The Nagars show trends towards retention, the lit. – illit. show trends towards non–retention. All the three groups are significantly different along the scale of retention of ã : illit. – lit. – Nagar.
–ã– and – ã

(j) The difference is significant at 1 percent level of confidence among lit. and at 5 percent level of confidence among illit. (confidence level among Nagars not counted).

The word final – ã– and the utterance final – ã appear along the same scale of retention of the nasal ã : the trends among the illiterates and the literates are towards non–retention and among the Nagars are towards retention. The behaviour of the illiterate and the literate groups is, as elsewhere, significantly different for the word final and the utterance final ã. It indicates that the position of the nasal has a significance – juncture is also a relevant unit in stating this variation.

2. –ũ– (tables 14–21) :

URBAN (MIXED)

[–ũ–]

(14)

Illiterate

Speakers	Frequency
5	1
17	2
16	3
10	4
0	5
1	6
0	7
1	8
—	
50	

Mean	—	2.82
Variance	—	1.5876
ns^2	—	79.38
$\frac{ns^2}{n-1}$	—	1.62

Population : 50
Frequency : 11

(15)

Literate

Speakers	Frequency
zero	1
9	2
8	3
11	4
6	5
8	6
2	7
1	8
3	9
1	10
1	11
—	
50	

Mean	—	4.68
Variance	—	5.05776
ns^2	—	252.88
$\frac{ns^2}{n-1}$	—	5.1608

Population : 50
Frequency : 11

URBAN (WOMEN)

[-ũ-]

(16) Illiterate		(17) Literate		(18) Nagar	
Speakers	Frequency	Speakers	Frequency	Speakers	Frequency
4	zero	10	zero	1	1
7	1	7	1	7	2
9	2	8	2	4	3
5	3	—		8	4
—		25		5	5
25				—	
				25	

Mean	— 1.60
Variance	— 0.96
ns^2	— 24.0
$\frac{ns^2}{n-1}$	— 1.00

Population : 25
Frequency : 5

Mean	— 0.92
Variance	— 0.7136
ns^2	— 17.84
$\frac{ns^2}{n-1}$	— 0.7433

Population : 25
Frequency : 5

Mean	— 3.36
Variance	— 1.4304
ns^2	— 35.76
$\frac{ns^2}{n-1}$	— 1.4900

Population : 25
Frequency : 5

NON-URBAN (MIXED)

[-ũ-]

(19) Illiterate		(20) Literate		(21) Nagar	
Speakers	Frequency	Speakers	Frequency	Speakers	Frequency
19	zero	3	zero	1	zero
3	1	4	1	2	1
1	2	3	2	1	2
1	3	2	3	1	3
1	8	1	10	2	5
—		—		2	6
25		13		2	7
				4	8
				4	9
				5	10
				—	
				24	

Mean	— 0.64
Variance	— 2.7904
ns^2	— 69.76
$\frac{ns^2}{n-1}$	— 2.9067

Population : 25
Frequency : 10

Mean	— 2.00
Variance	— 6.3077
ns^2	— 32.0
$\frac{ns^2}{n-1}$	— 6-8333

Population : 13
Frequency : 10

Mean	—6.7083
Variance	—9.7904
ns^2	—234.9688
$\frac{ns^2}{n-1}$	—10.2160

Population : 24
Frequency : 10

(a) In the urban area -ũ- illit. shows low mean and low variance-definitely displays non-retention ; variant behaviour is almost non-existent. -ũ- lit. shows higher mean and higher variance, though majority are grouped in a low frequency range, and the tendency is at non-retention.

(b) The difference between lit. and illit. is significant in both, the nonurban and the urban areas. Lit. – illit. difference is significant at 1 per cent confidence level.

(c) The lit. -illit. women informants (urban) do not show any significant difference in -ũ- u treatment; both display low mean, low variance.

(d) The Nagars, urban (women) and non-urban differ significantly. In the urban areas they show high mean and low variance, which indicatea a definite trend towards retention; in the non-arban area they show high mean and high variance (retention, but indecision).

-ũ (tables 22-26):

URBAN (MIXED)

[-ũ]

(22)
Illiterate

Speakers	Frequency
2	0
1	1
6	2
14	3
6	4
11	5
1	6
6	7
1	8
0	9
1	10
1	11
50	

Mean	—	4.22
Variance	—	4.9316
ns^2	—	246.58
$\frac{ns^2}{n-1}$	—	5.0322

Population : 50
Frequency : 11

(23)
Literate

Speakers	Frequency
4	1
4	2
7	3
5	4
5	5
7	6
4	7
3	8
7	9
2	10
2	11
50	

Mean	—	5 54
Variance	—	8.0484
ns^2	—	402.42
$\frac{ns^2}{n-1}$	—	8.2127

Population : 50
Frequency : 11

URBAN (WOMEN)

[-ũ]

(24) Illiterate Speakers	Frequency	(25) Literate Speakers	Frequency	(26) Nagar Speakers	Frequency
7	zero	9	zero	1	zero
8	1	9	1	1	1
5	2	2	2	2	2
1	3	5	3	5	3
4	4	—		7	4
—		25		9	5
25				—	
				25	

Mean	— 1.48	Mean	— 1.12	Mean	— 3.72
Variance	— 1.8496	Variance	— 1.2256	Variance	— 1.8017
ns^2	— 46.24	ns^2	— 30.64	ns^2	— 45.04
$\frac{ns^2}{n-1}$	— 1.9267	$\frac{ns^2}{n-1}$	— 1.2767	$\frac{ns^2}{n-1}$	— 1.8767
Population : 25 Frequency : 5		Population : 25 Frequency : 5		Population : 25 Frequency : 5	

(e) In the urban areas, -ũ among the literates displays a deviant behaviour at retention non-retention levels while among the illiterates it displays a deviant behaviour of a different type : both the extremes of frequency are well distributed among the speakers but the majority is showing indecision more in favour of non-retention.
The difference in lit. -illit. behaviour is at 5 per cent level of confidence.

(f) Among the women informants the lit.-illit. difference is not significant.

(g) The Nagars distinctly behave towards retention, high mean and low variation.

–ũ– and –ũ ·

(h) The difference is significant at 5 percent level of confidence among the lit. and at 1 percent level of confidence among the illit. (confidence level among the Nagars not counted).

As in the case of ā, in the case of ū also, the position of the nasal vowel is significant and the behaviour of the illiterate and the literate groups is also significantly different though both these groups tend towards non-retention as against Nagars who tend towards retention of the nasal ũ.

3. The second variable is an initial d– alternating with an inital ḍ– when there is a retroflex ṭ, ḍ, or ḷ in the following syllable. There are about ten words in Gujarati with such a structure and all these words were included in the questionnaire; the samples were obtained from Ahmedabad (all women informants) and Surendranagar (mixed) from the three groups, illiterate, literate and Nagar. The items were :

daṭ–	'to bury'	daḷia	'roasted chick–peas'
dɔṭ	'run'	dīṭū	'stem, stalk'
dūṭi	'navel'	dɔḍh	'1 ½'
dāḍo	'a stick'	daḍhi	'chin, beard'
daḷ	'lentil'	daḍh	'molar teeth'

The spelling system retains d– in all these cases.

d– > ḍ– (tables 27–32) :

URBAN (WOMEN)

d– > ḍ –

(27) Illiterate		(28) Literate		(29) Nagar	
Speakers	Frequency	Speakers	Frequency	Speakers	Frequency
1	3	3	1	1	zero
2	4	1	2	3	1
3	5	2	3	4	2
4	6	5	4	6	3
2	7	1	5	5	4
3	8	4	6	4	5
6	9	1	7	1	6
4	10	4	8	1	7
—		2	9		
25		2	10	—	
		—		25	
		25			

Mean	— 7.28
Variance	— 4.4416
ns^2	— 111.04
$\frac{ns^2}{n-1}$	— 4.6267

Population : 25
Frequency : 10

Mean	— 5.48
Variance	— 7.6096
ns^2	— 190.24
$\frac{ns^2}{n-1}$	— 7.9267

Population : 25
Frequency : 10

Mean	— 3.28
Variance	— 2.7616
ns^2	— 69.04
$\frac{ns^2}{n-1}$	— 2.8767

Population : 25
Frequency :10

NON-URBAN (MIXED)

d- > ḍ-

(30) Illiterate		(31) Literate		(32) Nagar	
Speakers	Frequency	Speakers	Frequency	Speakers	Frequency
1	6	1	3	3	0
2	8	1	4	2	1
6	9	3	6	2	2
7	10	1	7	1	3
9	11	1	8	4	4
—		2	9	2	5
25		2	10	5	6
		2	11	1	7
		—		3	8
		13		1	11
				—	
				24	
Population : 25		Population : 13		Population : 24	
Fraquency : 11		Frequency : 11		Frequency : 11	

(a) There is a greater incidence of d > ḍ in the non-urban areas; in all the three groups; illit. lit. and Nagars.

(b) Thus, the urban illiterates (women) show about half the population in high frequency range, none in low frequency and a little less than half in the middle frequency range; the non-urban illiterates (mixed) show most of the speakers in high frequency range, and none in the low frequency range.

Both display a definite trend towards ḍ.

(c) The urban literates (women) show a low mean and high variance with about 1/3 speakers in the high frequency range, 1/4 in the low frequency range and 1/2 in the middle range; this indicates greater indecision and drift in the urban literates; the non-urban literates (mixed) show majority of the speakers in high frequency group.

(d) The urban Nagars (women) show low mean and low variance; there is none in the high frequency range, less than half in the middle frequency range and more than half in the low frequency range, this indicates a definite trend towards d, and rejection of d>ḍ. Among the non-urban Nagars (mixed) about 1/5 speakers are in the high frequency range, 1/5 in the low frequency range and the rest in the middle frequency range; it indicates greater indecision and drift (somewhat comparable to the urban literates: non-urban Nagars = urban literates).

The above two cases, the nasal vowels ĩ / ũ~i / u and d~ḍ are variations which involve a phonemic change (partial merger with zero and retroflex ḍ respectively) as the nasal vowels and the dental- retroflex contrasts are distinctive in Gujarati. In both these items, the spelling rules prescribe the marking of the nasalised vowels and retention of d in the writing system.

The third case is that of an allophonic variation- selection of the allophones of / ph / among the different social groups. In the case of / ph / there is only one symbol in the writing system.

Gujarati / ph / has three allophones: [ph], [f] and [ɸ]; [ph-] occurs initially; medially, all the three vary freely. A sample of twenty items with the medial allophones was obtained from Ahmedabad (urban, women) and Surendranagar (non-urban, mixed). In the sample study, the variation between [ɸ] and [f] was not separated in the non-urban area, both the allophones were transcribed by one symbol [f] whereas in the urban area the allophones were separately transcribed as [ɸ] and [f].

For the purposes of this study, initial ph- means utterance initial ph-, and medial -ph- means utterance medial -ph-. The questionnaire is, therefore, in two sections; the first ten questions include word initial ph-; both are interpreted as utterance medial -ph-. (It is apparent that a more refined questionnaire which would neatly separate all allophones- prejunctural and postjunctural-

could have been designed, and it could have yielded more refined result; apologies for this shortcoming).

Allophones of/ph/in non-initial position (tables 33-38) :

NON-URBAN (MIXED)

non-initial [ph]

(33) Illiterate		(34) Literate		(35) Nagar	
Speakers	Frequency	Speakers	Frequency	Speakers	Frequency
1	3	1	2	1	zero
1	4	1	4	2	1
1	5	1	8	1	3
2	6	2	9	2	5
1	7	1	10	1	6
1	8	3	11	1	7
1	9	1	12	3	8
5	10	2	13	4	9
4	11	1	17	5	10
4	12	—		3	11
3	13	13		1	13
1	15			—	
—				24	
25					

Mean — 9·8	Mean —10.00	Mean —7.6667
Variance — 9.12	Variance —13.8462	Variance—11.8050
ns^2 — 228.0	ns^2 —180.0	ns^2 —283.3211
$\frac{ns^2}{n-1}$ — 9.5	$\frac{ns^2}{n-1}$ —15.0	$\frac{ns^2}{n-1}$ —12.3183
Population : 25	Population : 13	Population : 24
Frequency : 20	Frequency : 20	Frequency : 20

URBAN (WOMEN)

non-initial [ph]

(36) Illiterate		(37) Literate		(38) Nagar	
Speakers	Frequency	Speakers	Frequency	Speakers	Frequency
4	14	1	7	1	0
2	15	1	9	2	5
2	16	1	12	3	7
3	17	1	15	2	8
3	18	2	16	4	9
3	19	4	17	3	10
7	20	2	18	2	11
—		8	19	4	12
24		5	20	1	13
		—		2	14
		25		1	18
				—	
				25	

Mean	— 17.5	Mean	—17.24	Mean	—9.68
Variance	— 4.9167	Variance	—10.9024	Variance	—12.3
ns^2	— 118.0	ns^2	—272.36	ns^2	—309.44
$\frac{ns^2}{n-1}$	— 5.1304	$\frac{ns^2}{n-1}$	—11.3567	$\frac{ns^2}{n-1}$	—12.8933
Population : 25		Population : 25		Population : 25	
Frequency : 20		Frequency : 20		Frequency : 20	

(a) It is noticed that the Nagars show greater indecision and drift : they use [f] or [ϕ] for more than half of the occurrences in the urban areas and for 3/5 of the occurrences in the nonurban area.

Below are given two counts of the total occurrences of [ϕ], [f] and [ph]; these counts are arrived at from the tables given above; on the basis of the occurrences of a given item in the questionnaire multiplied by the population, one can arrive at the total occurrences of that item. Thus, for example, of the total 500 occurrences of (25 population x

20 frequency), among the non-urban illiterates (table 33), [ph] occurs for 245 times and [f] for 248 times (the minor discrepancies in the total are due to doubtful entries in the data); similarly out of the total 260 occurrences (13 population x 20 frequency), among the non-urban literates (table 34), [ph] occurs for 130 times and [f] for 140 times (allowing for the minor discrepancy in the total).

In the non-urban area, the incidence of -ph- and -f- is

	-ph-	-f-
illit.	245	248
lit.	130	140
Nagar	184	288

which distinguishes the non-urban Nagars from the rest by a noticeable higher frequency of [f] (or [ϕ]) as against the lit. and illit. groups. Moreover, the range of variation in the use of [ph] is from zero to thirteen, and the frequencies in that range are well spread out, which indicates indecision.

(c) In the urban area, the incidence of -ph-, -f- and -ϕ-, is

	-ph-	-f-	-ϕ-
illit.	420	41	11
lit.	341	48	8
Nagar	242	121	131

This brings out the Nagar – nonNagar difference sharply; the non-Nagars display -ph- almost exclusively with very little variation while the Nagars display -f- or -ϕ- half the time.

Some clear patterns emerge when we correlate the occurrences of these variables among the various strata of the population. The urban – non-urban social situations are different and there is a sharper ranking in the urban scales :

non-urban illiterate vs. urban illiterate
vs.
non-urban literate vs. urban literate
vs.
non-urban Nagar vs. urban Nagar

One would, therefore, like to investigate further in this direction to find out the nature of social classes, which is probably more definitive, with discrete boundaries in the

urban areas. In both the cases, urban as well as non-urban, a single group seems to be distinct from the rest, probably serving as the arbiter of linguistic fashions. This group, the Nagars, focuses our attention on the social elite-consciousness of which literacy is probably merely one of the many dimensions, such as : occupation, income, authority-power, etc. The difference within the Nagars (urban versus non-urban) on the one hand is a part of the general difference between the urban- non-urban differences and on the other, it may stem from the 'regional differences' in social recognition; probably, it is only in the older capitals (historically, older seats of power) such as Ahmedabad, Junagarh and Mandvi (Kutch) that Nagar is recognized as the apex; hence the difference between Nagars of Ahmedabad and those of Surendranagar in our present inquiry.

In one variable, (nasalised ã) the behaviour of literate and illiterate women does not fit in the pattern, and seems to sugest that there may be some significant difference in the speech behaviour of men and women in our community. -ã- and -ã which generally show the tendency of losing their nasalisation along the illiterate – literate scale and retention of nasalisation among the Nagars, appear with a greater scatter and a high mean and variance of -ã- and -ã among illiterate women; though, the difference between illiterate and literatc women does not come within confidence level.

Before interpreting this deviant behaviour, it may be noticed that no such deviation occurs in u/ũ variation. It is probable that since in u/ũ variation the loss of nasalisation is significant by almost non-existant variant behaviour (at least in one case : -ũ- urban illiterate) the chances of its retention are considerably reduced among illiterate or literate women.

Some tentative explanations may be offered for this deviant behaviour of women. The illiterate women are surrounded by literate men in the family, particularly in urban households, hence they do not behave like illiterate

men. Urban illiterate women are, as women in literate contexts, half-way between literate women and illiterate men. Secondly, literate women in the sample happen to be mostly elite castes and families – not true of the literate men as such in an urban area. The literate woman is by definition high caste, if not the highest; on the other hand, even the illiterate woman in urban area has intimate literate company of brother, husband and sons and daughters and so on. These factors may be taken into consideration in interpreting our samples. Lastly, I should take into account the bias of the investigator's universe – a woman – in this case.

The limitations of this study are apparent; the samples in different regions and different strata of population have not been evenly worked out; the samples are not properly refined i. e. in the nasalisation study, I have worked with only literate–illiterate distinction in one group; each could have been broken down castewise and sexwise. The samples have been quite random. I have not taken into account age differences, nor have I taken into account any other scales of social stratification – income, residence, occupation, etc. Moreover, this fieldwork was done before 1962 and before I had a chance to study some excellent work in sociolinguistics – Labov's studies in New York English; I have not used Labov's style indices.

In spite of these shortcomings, this study may definitively show that simple models of correlations between caste ranking and speech variation cannot explain the complex structure of variation; the parameters are many and complex; the developing urban centres display emerging social classss with considerable vertical mobility-which indicate that variation is functional, variation does, not indicate a barrier in communication. Probably a study of speech variation in a country like India which is rapidly moving away from a traditional society to a modern society may throw light on the formation of social classes, ancient as well as modern.

APPENDIX TO LECTURE 4 : THE QUESTIONNAIRE PARAMETERS OF SPEECH VARIATION IN AN INDIAN COMMUNITY

The Questionnaire is in Gujarati and it is reproduced below in Nāgarī transcription. The fieldworker kept the questionnaire with him and spoke out the sentence, asking the respondents to repeat the sentence in normal casual style. The respondents did not know what was intended, sometimes they inquired and were told that this is a survey of how people speak the Gujarati language.

1. QUESTIONNAIRE

URBAN (MIXED)

[ũ] and [ã] in different positions with different types of consonants.

1. एक साथे वार पोलकां ?
2. बोर पाकां लावजो.
3. तमे जरा वार नाका बहार बेसो.
4. रामजीभाई भारे पक्का.
5. वनिताबेन तो बहु ज पोचां.
6. ए साचां बहेनो छे.
7. एमनी वाचा हराई गई छे .
8. बधा ज एवा लुच्चा ?
9. केम आप्यां टमेटां ?
10. अमे खाटां फळ लाव्यां.
11. रेलना पाटा पासे थी.
12. तारा जेवा घणा जोया, भामटा ?
13. खूब धारदार छे त्रणे चप्पां.
14. कोई छापां एवी जाहेरखवर न ले.
15. धूळेटी पडवे थापा तो मारे ज.
16. अररर ! एक साथे त्रण कापा ?
17. बेबीबेन सवारे गीत गातां.

18. पाणी पीतां वार केटली ?
19. उनाळे वायरा बहु वाता.
20. तमारा पिता आजे आवशे ?
21. अरर्र ! आम भराय वचकुं ?
22. बराबर वच्चे ओशीकुं मूको.
23. मने थोडी तमाकु आपो.
24. कातर नहीं, चाकु.
25. ए तो मरचुं छे मरचुं .
26. तमे साचुं कही दो पछी आगळ वधाय.
27. भरोसे न रहेशो. हचुडचु छे.
28. ए तो लालचु छे लालचु.
29. हजी भावनगर केटलुं छेटुं ?
30. ए मोटुं काम छे.
31. पछी पाटु मारीश.
32. तो करो साटु.
33. तमारी साथे कोण रहेतुं ?
34. बहार कोइक गातुं जाय छे.
35. तू तू करो तो कूतरो आवे.
36. पोलाद तो बहु मजबुत धातु.
37. केरी अत्यारे कापुं ?
38. ना छापुं जोया पछी.
39. मने चप्पु आपो.
40. आपशे तारा बापु.
41. पछी काकी खीजायां.
42. लीलीबेन तो सहु थी सवायां छे.
43. तमे लोभाया वगर बोलो.
44. ए बोल्या वगर गया.
45. कोण आव्युं?
46. हवे थयुं, छोडो वात.
47. तेनो पियु परदेश छे.
48. आ वायु तो जो वायु.
49. छे य आपणी बेबी नी जेम गोरां.
50. लीलाबेन सारां लागे छे.
51. जाडा दोरा वापरजो.

52. बधा ज चोपडा कोरा.
53. ए ज घर तमारूं.
54. तमे सारूं न लगाडो.
55. भले; वारू, तमे कही तेम.
56. ए वाघ छे के वरू ?
57. कमुबेन काले आवजो वहेलां.
58. जरा लीलां पान लावजो.
59. गाडीना चीला जोता रहेजो.
60. साहेब नकामा बहु तपेला.
61. चहेरो जोईने थाय टीलुं.
62. कोईक ज भलुं करे छे.
63. मोटर चालु करीने बोलावो.
64. च्हा बनावोने नीलु.
65. ए परणे तो गीत गावां.
66. बा केवां देखाय छे ?
67. सेवा करो तो मेवा मळे.
68. जलदी खावा दे खावा.
69. एक पहेरण सीवुं ?
70. मारे पाणी पीवुं छे.
71. पेलो देवु कोण छे ?
72. देवु, ए देवु !
73. बिचाराना पाप शां ?
74. तेने कशांनी समज पडती नथी.
75. रोज नशा थी चकचूर होय छे.
76. फोडी दो ने शीशा.
77. पडशे तेवा देशुं.
78. अमे कशुं न जाणीए.
79. तो जशु जाणे छे ?
80. साव पशु छे पशु.
81. पहेरण ने य चार खीसां ?
82. मासी तो खासां कलाक थी छे !
83. साथे मासा पण छे !
84. अमे तो वीसा.
85. तमे कहो तो हसुं.

86. साबु घसुं पछी आवो.
87. मारी सासु खीजाशे.
88. घटे छे एकाद तसु.
89. पोलिस जोईने छू थई गयो.
90. दाढीए जू वळगी छे.
91. चार किलो रू बस छे.
92. हवे लू शरू थवानी.
93. ए पाणी ने भू नथी कहेतो.
94. पेलो वाटको क्यां ?
95. आजे हुं घेर ज रहीश, तुं ?
96. काले तुं बहार जशे के हुं ?
97. एटले शुं तने भाररूप छुं ?
98. भले छुं तो घेर, पण तेथी शुं ?
99. तमे त्यां क्यां जाओ छो ?
100. तमे ज्यां कहो त्यां ?

2. Questionnaire :— Urban – Women

Non – final [-ũ-]

1. आम बचकुं मराय ?
2. ना. छापुं जोया पछी.
3. सवारे बधुं पूछजो.
4. कोईक ज भलुं करे छे.
5. ए मोटुं काम छे.

Non – final [-ã-]

1. पाणी पीतां वार केटली ?
2. कोई छापां छे ?
3. जरा लीलां पान लावजो.
4. आ तमारां कपडां.
5. कपडां तो दूध जेवां ऊजळां.

Final [-ũ]

1. अत्यारे केरी कापुं ?
2. तमे कहो तो हसुं.
3. हुं तो परोढिये ज ऊठुं.
4. आ तमारूं ओशिकुं ?
5. जरा जोर थी दाबुं ?

Final [-ã]

1. एक साथे बार पोलकां !
2. बेबीबेन सवारे गीत गातां.
3. वनिताबेन तो बहु पोचां.
4. पछी काकी खीजायां.
5. सविताबेन सौथी सवायां.

3. Questionnaire : Non – Urban (mixed)

Non - final [-ũ-]

1. हुं हसुं छुं.
   अरे हुं ते हसुं ?
2. हुं कापुं छुं.
   हुं ते कापुं ?
3. हुं सवारे वहेलो ऊठुं छुं.
   काले हुं केटला वागे ऊठुं ?
4. मने जाडुं ओशीकुं फावतुं नथी.
   आ ज ओशीकुं ?
5. आटलुं शाक तो घणुं छे.
6. कूतराए बच्चकुं भर्युं !
7. तमे सवारनुं छापुं वांच्युं ?
8. आटलुं बधुं बोलवानुं कोणे कह्युं ?
9. भगवान तमारूं भलुं करो.
10. आवी बाबतो मां तमारे मोटुं पेट राखवुं जोईए.

Non—final [ã]

1. मारां कपडां केटलां ऊजळां छे !
2. वेन मारा ऊपर खीजायां हतां.
3. आजेतो छापांनो ढगलो थ्यो.
4. तमने लीलां मरचां फावे ?
5. केटलाक छोकरा छोकरीयो गातां हतां.
6. वडां पोचां छे.
7. आ छोकरां तमाराथी सवायां थशे.
8. तमे पाणी पीतां पीतां वात न करो.
9. तमने तो सारीरीते चालतांय नथी आवडतुं.
10. तमे जरा बोलतां तो शीखो !

4. Questionnaire : Urban—women [d ~ ḍ]

1. साव दाट वाळी नाख्यो.
2. हाथमां दंडो पकडयो छे.
3. जो, केवी दोट मूकी !
4. उपरनो दांडो पकडी लो.
5. केलेण्डर मां दट्टो भराव्यो छे.
6. नवी दाढ कयारे आवी ?
7. आजे दाळ सरस थई छे.
8. अत्यारे दोढ वाग्यो छे.
9. उपरथी दीट्टुं काढी नाखो.
10. कूतरूं करडे तो दूंटीनी आजुवाजु इन्जेकशन लेवां पडे.

5. Questionnaire : Non—Urban mixed

[d ~ ḍ]

1. डोशी ए तो दाट वाळयो.
2. एना माथा मां मोटो दंडो वाग्यो.
3. नानां छोकरां ने दाळिया बहु भावे.
4. एणे एवी तो दोट मूकी !

5. नर्मद दांडियो नामनुं छापुं काढतो हतो.
6. आ कॅलेण्डर नो दट्टो खलास थयो छे.
7. मारी दाढ सडे छे.
8. मने लींबु नीचोवेली दाळ फावती नथी.
9. केटलाक माणसो दोढडाह्या होय छे.
10. रींगणां नां दींटियां रहेवा देजो.
11. कुतरूं करडे तो दुंटीनी पासे इन्जेकशन लेवुं पडे.

6. Questionnaire : Urban (women) and Non—Urban (mixed).
word-medial [– ph – ~ – ɸ –]

1. खोटी अफवाओ सांभळवी नहीं.
2. कोई कंई मफत ओछुं ज आपे ?
3. आज घणो बफारो छे.
4. तमारा दफतर मां चोपडीओ छे ?
5. तो पछी सारा-नरसा वच्चे शो तफावत ?
6. आ रूम नी सफाई सारी छे.
7. तमे रफते रफते आ काम शीखी जशो.
8. केटलाक माणसो डफोळ होय छे.
9. मने कंई पैसादारो माटे नफरत नथी.
10. क्रीडीओनो राफडो फाटयो.

7. Questionnaire
Urban (women) and Non-Urban (mixed)
word initial [ph – ~ ɸ –]

1. एटलामां तो फडाफडी बोली गई.
2. दीवाळी मां फटाकडा फोडवानी मजा आवे.
3. कोलेजियनो फककड थईने फरे छे.
4. एटली नानी वात ऊपर रजपूतो फना थता.
5. एमां शुं ? ए तो मारी फरज छे.
6. केरी जेवुं उत्तम फळ बीजुं नथी.
7. ठाकरडा माथे फाळियुं बांधे.
8. आ छोकरा फाट्या छे हवे.
9. फाळका ऊपर सूतर वींटो.
10. तमारी फई नुं नाम शुं ?

REFERENCES

AGESTHIALINGOM, S. 1967 — A Generative Grammar of Tamil. Annamalai University, pp. 55–73.

ALLEN, W. S. 1951 — *Phonetics in Ancient India*, London.

BANATHY, BELA, EDITH TRAGER, AND CARL D. WADDLE 1966 — The Use of Contrastive Data in Foreign Language Course Development. In A. Valdman (Ed.), *Trends in Language Teaching*, New York: McGraw Hill. pp. 35–56.

BLOCH, JULES 1920 — Formation de la langue Marathe. Paris.

BLOOMFIELD, LEONARD 1933 — *Language*. New York: Holt.

BRIGHT, WILLIAM 1960a — Linguistic Change in Some Indian Caste Dialects. In Charles A. Ferguson and John J. Gumperz (Eds.) *Linguistic Diversity in South Asia*, supplement to IJAL, 26. 19–26.

1960b — Social Dialect and Language History. *In Current Anthropology*, 1.424–5.

CHATTERJI, SUNITI KUMAR 1926 — Origin and Development of the Bengali Language, Calcutta.

DIMOCK, EDWARD C. 1960 — Literary and Colloquial Bengali in Modern Bengali Prose. In *Linguistic Diversity in South Asia*, supplement to IJAL, 26.43–63.

EMENEAU, M. B. 1956 — India as a Linguistic Area. Lg. 32.3–16.

ERVIN-TRIPP, SUSAN 1967 Commentary on R. M. Jones's paper, How and When Do Persons Become Bilingual ? In L. G. Kelly (Ed.) *Description and Measurement of Bilingualism* pp. 26–35. Toronto.

FERGUSON, CHARLES 1959 Diglossia, Word 15.325–340; reprinted in Del Hymes (Ed.) *Language, Culture and Society*. New York : Harper and Row. 1964, pp. 423–440

FERGUSON, CHARLES AND JOHN GUMPERZ 1960 (Eds.) *Linguistic Diversity in South Asia*, supplement to IJAL, 26.

GRIERSON, GEORGE ABRAHAM 1908 *Linguistic Survey of India*, Vol. IX, Part II, pp. 447–8, Calcutta.

GUMPERZ, JOHN J. 1958a Dialect Differences and Social stratification in a North Indian vollage. *In American Anthropologist*, 60.668–82.

1958b Phonological Differences in three Hindi Dialects. Language. 34.212–24.

1960 Formal and Informal Standards in the Hindi Regional Language Area. In Charles A. Ferguson and John J. Gumperz (Eds.) *Linguistic Diversity in South Asia*, supplement to IJAL. 26. 92–118.

1964a Hindi-Punjabi code-switching in Delhi. In Horace G. Lunt (Ed.) *Proceedings of the 9th International Congress of Linguists.* The Hague : Mouton, Pp. 1115–1124.

1964b Linguistic and Social Interaction in Two Communities. In John J. Gumperz and Dell Hymes. *The Ethnography of Communication.* American Anthropologist. 66 (6), part 2.137–153.

1966 On the Ethnology of Linguistic Change. In William Bright (Ed.) Sociolinguistics. The Hague : Mouton, 1966. 27–49.

1969a ' How can we Describe and Measure the Behaviour of Bilingual groupsa ' In L. G. Kelley (Ed) *Description and Measurement of Bilingualism* pp. 241–284, Toronto.

1969b ' Sociolinguistics in South Asia ' In Thomas A. Sebeok (Ed) *Current Trends in Linguistics–Vol.* 5 : *Linguistics in South Asia,* Mouton, 597–606.

HAUGEN, EINAR 1956 *Bilingualism in the Americas.* A Bibliography, and Research Guide, American Dialect Society, University of Alabama Press, Alabama.

KAKATI, BANIKANTA 1941 *Assamese, Its Formation and Devlopment* Gauhati.

KELKAR, ASHOK R. 1968 *Studies in Hindi–Urdu.* Deccan College, Poona.

LABOV, WILLIAM 1967 Commentary on J. J. Gumperz's paper, How can We Describe and Measure the Behaviour of Bilingual Groups ? In L. G. Kelly (Ed.) *Description and Measurement of Bilingualism.* pp. 250–55. Toronto.

MACKEY, WILLIAM F. 1965 Bilingual Interterence : Its Analysis and Measurement. Journal of Communication, 15. 235–249.

MACNAMARA, JOHN 1967 How can one Measure the Extent of a Person's Bilingual Proficientcy ? In L. G. Kelly (Ed.) *Description and Measurment of Bilingualism,* pp. 80–97. Toronto.

MALHERBE, E. G 1967 Commentary on R. M. Jones's paper, How and When Do Persons Become Bilingual ? In L. G. Kelly (Ed.) *Description and Measurement of Bilingualism,* pp. 41–51. Toronto.

OSGOOD, CHAREES E. AND THOMAS A SEBEOK
1965 *Psycholinguistics: a Survey of Theory and Research Problems,* Indiana University Press.

PANDIT, P. B
1961 Historical Phonology of Gujarati Vowels, *Language,* 37. 54–66.

1963 'Sanskritic Clusters and Caste Dialects' *Indian Linguistics*, Vol. 24, 1963. 70–79.

1969a 'Parameters of Speech Variation in an Indian Community' In 'Language and Society in India' Transactions of Indian Institute of Advanced Study, Vol. 8. Simla Pp. 207–229.

1969b Comments on J. J. Gumperz's paper, 'How can we Describe and Measure the Behaviour of Bilingual Groups?' In L. G. Kelley (Ed) Description and Measurement of Bilingualism, Toronto, Pp. 255–256.

1968 A Grammar of Gujarati Number Names, Indian Linguistics, Vol.29, Part I. Pp.132-41.

PILLAI, M. SHANMUGAM
1960 Tamil–Literary and Colloquial. In Charles A. Ferguson and John J. Gumperz (Eds.) *Linguistic Diversity in South Asia,* Supplement to IJAL, 26. 27–42.

PRZYULSKI, JEAN (*Ed.*) *Pre–Aryan and Pre–Dravidian.*

RADHAKRISHNAN, R.
1963 'Tamil Numerals' Grammar', Indian Linguistics 24 (1963) Pp. 110–122.

RAMSUBRAMANIAN, K. S.
1968 *A Linguistic Study of the Saurashtri Language.* Unpublished dissertation, Poona University.

RANDLE, H. N.
1943–46 'An IndoAryan Language of South India: Saurashtra Bhāṣā' BSOAS XI. pp. 104–121 and 310–327.

SAKSENA, BABURAM 1937 *Evolution of Awadhi*, Allahabad.

SAPIR, EDWARD 1939 *Language*, New York.

SHASTRI, MANGALA DEVA 1931 (Ed.), *Rk Prātiśākhya*, Allahabad.

SPENCER, JOHN 1957 Received Pronunciation, Some Problems o Interpretation, *Lingua* 7, 7–29.

TURNER, RALPH L. 1931 *A Comparative and Etymological Dicti onary of the Nepali Language, London.*

VAN KATWIJK, A. 1965 A Grammar of Dutch Number Names *Foundations of Language*, 1. 51–58.

WEINREICH, URIEL 1967[5] *Languages in Contact*, Mouton.