

CHAPTER - II

MATERIALS AND METHODS

CHAPTER - II

Materials and Methods

2.1 Materials - American, British and Indian Corpora

Our dissertation is based on the occurrences of weak forms in the Kolhapur corpus of Indian English (Shastri etal - 1986) A short description of the corpus is given below.

Katkar (1984) has given a detailed description of the history, rationale and procedures adopted in the building of Kolhapur corpus of Indian English in her Ph.D. dissertation. And it has been convincingly argued that the corpus is adequate as source material for studies of different aspects of I.E. Here, we will deal with some important and necessary details of the corpus.

America was the first country to build 'A Standard Corpus of Present Day Edited American English, (Francis etal- 1979) It provides a standard data base of American English for the purpose of linguistic studies of American English. It also serves as a model for similar corpora of other varieties of English to be built. Within the ten years of the building of the American corpus (Known as Brown corpus) many important studies of different aspects of American English have appeared.

For Example - Componentational Analysis of Present Day American English, Kurcera and Francis (1967)
The Meanings of the Modals in Present Day American English Ehrman (1966) and
A Statistical Study of the Graphic System of Present Day American English Zetterstein (1969)

During the same period a similar corpus of British English known as 'Lancaster Oslo/ Bergen' (LOB for short) was built by Leech and Johnson et al (1978). When this became available several studies of British English and also comparative studies of British and American English started appearing. The significant ones are -

'The Semantics of the Modal Auxiliaries, Coates (1983)
and Word Frequencies in British and American English, Hofland and Johnson (1982)

The Kolhapur corpus of Indian English was completed in 1986, which has been built on the same lines as the other two (Brown and LOB) Corpora. Like the other two it is a million word corpus for use on digital computers. It consists of 500 texts of 2000 words each. The texts are distributed over fifteen different categories of material which represents a whole variety of written styles.

Basically, there are two prose included in it.

- 1) Informative Prose (374 texts) and
- 2) Imaginative Prose (126 texts).

Shastri describes (1980,1985,1986) in detail the extent of comparability and the minor differences between the three corpora. The main and the obvious difference is that the LOB and Brown corpora draw their samples from materials published in 1978. The sampling procedures, the types of different styles of materials represented and the weightage given to different categories of materials have all been kept very close to the others. The basic composition of materials in the three corpora is given in Table No. 1.

On one hand table No. 1. gives the major difference between the LOB and the Brown corpora and on the other, Kolhapur corpus with its internal variation of weightage to different categories of imaginative prose (category 'K' to 'R')

Table No.1 - The Basic Composition of America, British and Indian English Corpora.

Text Categories	No.of text in each category		
	American corpus	British corpus	Indian corpus
A) Press : reportage	44	44	44
B) Press : editorial	27	27	27
C) Press : reviews	17	17	17
D) Religion	17	17	17
E) Skills, Trade and Hobbies	36	38	38
F) Popular Lore	48	44	44
G) Belles Letters	75	77	70
H) Miscellaneous (Govt. Documents, Foundation reports, industry reports, college catalogue, industry house organ)	30	30	37
J) Learned and Scientific Writings	80	80	80
K) General fiction	29	29	58
L) Mystery and Detective fiction	24	24	24
M) Science Fiction	06	06	02
N) Adventure and Western fiction	29	29	15
P) Romance and love story	29	29	18
R) Humour	09	09	09
Total	500	500	500

2.2 ICE TEXT CATEGORIES

2.2.1 SPOKEN - (300)

I) Dialogue (180)

a) Private (100)

Direct conversations (90)

Distanced conversation (10)

b) Public (80)

Class lessons (20)

Broadcast Discussions (20)

Broadcast Interviews (10)

Parliamentary Debates (10)

Legal Cross - Examination (10)

Business Transactions (10)

II) Monologue (120)

a) Unscripted (70)

Spontaneous Commentaries (20)

Unscripted Speeches (30)

Demonstrations (10)

Legal Representations (10)

b) Scripted (50)

Broadcast News (20)

Broadcast talks (20)

Speeches (not broadcast) (10)

2.2.2 Written - (200)

I) Non- Printed (50)

- a) Non- Professional Writing (20)
 - Student Untimed Essays (10)
 - Student Examination Essays (10)
- b) Correspondence (30)
 - Social Letters (15)
 - Business Letters (15)

II) Printed (150)

- a) Informational (learned) (40)
 - Humanities (10)
 - Social Sciences (10)
 - Natural Sciences (10)
 - Technology (10)
- b) Informational (popular) (40)
 - Humanities (10)
 - Social Sciences (10)
 - Natural Sciences (10)
 - Technology (10)
- c) Informational (reportage) (20)
 - Press News Reports (20)
- d) Instructional (20)
 - Administrative/Regulatory (10)
 - Skills/ Hobbies (10)
- e) Persuasive (10)
 - Press Editorials (10)
- f) Creative (20)
 - Novels/Stories (20)

2.3 - Description of selected Texts

We have selected 'Five' texts of different categories from I. C.E. Texts given above. We have restricted ourselves to the study of the pronunciation of the weak forms by the speakers. The following is the description of Texts.

I) T.V. News - (scripted) S2B 020

Wed, 22.5.91 by Minu (Female) This text is scripted monologue, which is broadcast news.

II) T. V. Talk (scripted) S2B 029

Sat, 17.11.90, by Prof. Manubhai Shah (age-52, Male) EMRC Gujarat
This text is scripted monologue which is broadcast talk

III) Class Lesson (public), (unscripted speech) - Non B C by
S.K. Verma. - S2A-034

It is a dialogue category, which is unscripted public dialogue.

IV) Telephonic Conversation - Dialogue - distanced conversation
S I A 092

A Dialogue between Mrs. Dessa and Mrs Nalini Date 27/08/92

V) Face to Face Conversation - Dialogue - Direct conversation
SIA - 068

Between Vina Menezes and Charul Jain

All the texts that we have selected have a variety. Two categories of texts are scripted monologues. Example. T.V. News, and T.V. Talk, where as other three categories are unscripted dialogues. Example. Class Lessons, Telephonic

Conversation and Face to Face Conversation. Again there is a variety in the selection of texts. The use of weak forms in monologue differs from dialogue, as well as it differs in scripted speech from that of unscripted speech. So we have located the use of weak forms in various kinds of situations.

2.4 - Methods

As we have mentioned earlier, we have selected 5 texts having variety, which are different from each other. The method of our study can be mentioned as below

I) Listening - All the weak forms have been carefully listened, the specific phonetic situation and the use, whether it is near R.P. or whether it is different from R.P. have been pointed out. The strong forms are used by Indian speakers. We have also located some example where the Indians tend to use strong forms.

II) Transcription - After listening all the weak forms, we have transcribed them in phonemically. While transcribing them, we have once again, listened to them. The transcription done by us is phonemic transcription. But we want to assert that some Indian phenomenon remains the same.

For Example - 1) Indians use /e/ instead of /ə/.

2) 'The' is pronounced as /d_nə/ and not /θə/.

3) /θ/ is pronounced as /t_n^h/.

We have not focussed these things in our dissertation but with reference to the transcription we have stated the use of weak forms by Indian English speakers, whether the Indian speakers knew about weak forms. A comparative study is made by us.

III) Statistical Measures - While describing every text, we have used statistics. We have made tables. The pronunciations of speakers have been divided in two types.

1) Which goes near to R.P. 2) Which is quite different from R.P. All the occurrences of weak forms are counted and percentages given.

V) Analysis of Weak Forms and Comments on the Use

After giving the tables, and the description of weak forms, we have analysed the different types of use of weak forms in different types of situations; we have also made some comments regarding the specific use of weak forms. We have also mentioned Indian variants which are used in place of R.P. Weak form with this method we have studied very minutely the pronunciation of weak forms by Indian speakers.