

CHAPTER THREE

PROCEDURE OF INVESTIGATION

3.1 Methodology

The methodology adopted in the investigation of prosodic phenomena in broadcast interviews basically consists of the following elements:

- (i) Observation, comparison and selection of data
- (ii) Transcription of data and in-depth examination of transcribed data
- (iii) Classification and systemisation of the data
- (iv) Generalisations based on what is observed

3.2 Theoretical Framework

The theoretical framework for this study, which is discussed in detail in Chapter 4, is based on a broad definition of prosody, i.e. that prosody subsumes the following auditory effects of speech: pitch, loudness, tempo and pause.

3.3 Description of Data

The data used in the study comprises recordings of approximately seven hours of broadcast interviews recorded over a period of four months (see Table 1, Appendix A for the details of the interviews recorded). The primary data consists of live recordings of broadcast interviews; thus whatever was presented during the interview was not edited or revised. Each interview is made up of two participants, the interviewer (marked A in the data) and the interviewee (marked B in the data). The duration of each interview is between 30 and 75 minutes. Lengthy interviews are necessary because they give the researcher sufficient speech material to examine how speakers segment both short and lengthy discourse, the presentation of information by both the interviewer and interviewee and the organisation of the interview talk.

The total number of participants in the interviews is nine and they are all Malays. The choice of Malay speakers is essential because the study seeks to examine the language of the native speakers of Malay. All the participants, with one exception, are males. They are all educated native speakers of Malay between the ages of 45 and 53.

Crystal and Davy (1975, p. 86) define conversation

... as any stretch of continuous speech between two or more people within audible range of each other who have the mutual intention to communicate, and bounded by the separation of all participants for an extended period.

Based on this definition one can thus categorise broadcast interviews as a kind of conversation. What distinguishes it from other forms of conversation is the number and roles of participants involved, degree of seriousness and formality and type of setting.

The broadcast interview is an interactional activity of considerable importance. It is a dialogic encounter which is essentially transactional. The participants enter such an encounter with the purpose of achieving some kind of goals in terms of the world outside the encounter. The interviewer's goal is to post questions which will elicit the required information and the interviewee's goal is to supply the information requested (cf. Asmah, 1995). The structure of the discourse which is characterised by sequences of question and answer is determined by these preestablished goals. It provides the audience with the feeling that they are part of the interview. Although the interaction is between the interviewer and interviewee, the prime audience is the listening-in audience.

Unlike informal conversation, participants of broadcast interviews know beforehand what they are going to talk about. The interview is planned according to a set of structures which is appropriate for interviewing and which is expected by the participant to be in operation for the duration of the interview. The spoken material itself is unscripted and the talk produced during the interview is spontaneous and unplanned. Thus the interviews recorded contain features of spontaneous speech such as hesitations, slips of the tongue, incomplete sentences, repairs and so forth. There is also evidence of English-Malay code-switching. The presence of English words in the speech of both the interviewer and interviewee (although both are native speakers of Malay) and features of unplanned speech indicate that the conversational interactions between participants are spontaneous and somewhat informal.

In order to attain a reasonable linguistic homogeneity, the speakers whose utterances were analysed were all chosen from a single dialect which is referred to here as standard spoken Malay, the kind of Malay used by the mass media, at schools, etc. and is geographically unmarked. The term "standard" is used to indicate more clearly the boundaries of the corpus, i.e. to draw a distinction between the kind of Malay being analysed and the regional dialectisms which would only be used and understood by the Malays from a particular state.

The subject matter brought up during the interview was not only relevant to the interviewee's line of work, field of specialisation or area of interest but was for the benefit of the audience who were listening-in.

The data is limited in style. It is cooperative adult speech in a particular setting (i.e. a recording studio), which is for the most part, goal-directed. Participants behave in a pleasant and cooperative manner.

3.4 Limitations

There were a number of restrictions imposed on this study. Therefore it does not aim to give an entirely representative picture of the speech of speakers of standard Malay, but rather present a description of the prosodic features realised in the speech of Malay speakers in a particular event, i.e. radio broadcast interviews. Because of the nature of the study itself which requires a detailed and thorough analysis of the data collected, the amount of data collected and analysed had to be limited to manageable proportions. The interference of noise which can affect the audibility of such prosodic features as loudness and pitch is another reason why recorded data is preferable to naturally occurring data (see also section 1.4 of this thesis).

Although the data is made of broadcast interviews, the present investigation would be regarded as a contribution to the study of Malay prosody rather than to the study of the prosody of interview language. In an age of mass-media like radio and television broadcast speech ought to be a worthy object of study in its own right.

3.5 Research Techniques

The study of Malay prosody necessitates application of research techniques that fit the purpose of reliably handling the research problem set out in Chapter One. The present approach aims at overcoming the drawbacks that are apparently inherent in the previous works on Malay prosody. As a whole, the important difference between the former and the latter is that due importance is given to the necessary distinction between mere guesses (derived from intuition, introspection, etc.) and instrumentally verified observations.

A point that has to be raised in any discussion of procedures is whether the study would benefit most from either auditory or instrumental technique, or whether a combination of both is required. Since this study is concerned with the investigation of prosodic phenomena at discourse level, it is felt that the corpus of data should be analysed auditorily first for such

an examination gives the researcher valuable information about what the listener hears or responds to.

The dangers of relying totally on instrumental analysis are highlighted by Danes (1960) as follows:

"Granting that instruments are more accurate and sensitive than the human ear, it is also true that instrumental records and their interpretations in terms of physical acoustics do not give us a true picture of the way in which the speakers hear and understand (evaluate) their own language. The significance and function of the various waves, formants, etc.- these must first be discovered, if only in outline, by an auditory analysis [sic] of spoken language".
(Danes 1960: 37)

Even Bruce and Touati (1990:44) who undertake to do an acoustic-phonetic analysis of dialogue prosody begins with an auditory analysis of their data for they consider "the auditory analysis in terms of a prosodic transcription" as "a useful categorisation of the basic prosodic structure of the utterances of dialogue". As regards instrumental analysis, theoretical objections often refer to the discrepancies between what is measured and what is heard. And since people rely solely on their ears in communicating with each other through speech, many researchers have been inclined to distrust the instrument and to base their research by means of auditory analysis (Crystal, 1969; Lindstrom, 1976).

Nevertheless, to lend support to the observations made by means of auditory technique, some of these auditorily analysed data were instrumentally analysed using Mac Speech Lab II which gave a display of Fo contour and measurements of duration. Mac Speech Lab can record, save and play speech with sampling rates from 5 to 80 kHz. and uses 12 bit resolution. The instrumental findings are intended to act as companion to the auditory observations. Beside Mac Speech Lab II, the researcher also used a sound card which was installed in an IBM compatible personal computer to check again the duration of pauses, fillers as well as duration of selected utterances.

3.6 Presentation of Data

The data of broadcast interviews are presented in terms of prosodic-oriented transcription, which is basically an orthographic transcription of selected fragments of the total data. To this orthographic transcription are added prosodic features selected from our descriptive framework presented in Chapter Four. Although the written transcription may not do justice to the wealth of information that is available to the listener in the original tape recorded data, an attempt was made to give as accurate a representation as possible of the original data from which it was derived. Thus any omissions of segments or syllables, hesitations, colloquial forms, repetitions, etc. are indicated in the transcription.

3.7 Notational Conventions

In the course of this study different sections of the data were used. Since the purpose of the analysis is to investigate the part that prosody plays in or within the domain identified for examination, different presentation procedures were followed as convenient. The data extracts in each chapter will be numbered sequentially, the chapter number preceding the number assigned to the data. Thus the first piece of data in Chapter Six will be designated Extract 6.1. The interview participants are identified by a letter, i.e. letter "A" identifies the speaker as the interviewer and letter "B" as the interviewee. The number placed before letters "A" or "B" indicates sequential arrangement.

The relevant conventions adopted in this chapter are as follows:

- (1) Recognisable 'words' uttered are transcribed using the conventional spelling of the items, regardless of the pronunciation. A phonetic transcription is avoided.
- (2) A certain degree of liberty is taken in transcribing certain brief responses in the data which are in the form of noises as mhm, oh, ah, mm, etc.

- (3) Relevant prosodic features are marked according to the conventions established in Chapter Four of this thesis.
- (4) The transcriptions generally follow the conventions established in the Conversational Analysis literature.

The following features of the transcription should be noted:

- [] square brackets indicate overlapping talk; the left hand bracket marks the beginning of the overlap, the right-hand bracket marks the end.
- ::: Colons indicate a lengthening of the sound just preceding them, proportional to the number of colons.

three dots indicate an incomplete utterance.

a hyphen at the end of a preinterrupted talk indicates that the continuation of the talk comes after the interruption.

a hyphen at the beginning of a post-interrupted talk indicates a continuation of the pre-interrupted talk.

- (0.45) Numbers in parenthesis indicate durations of silence, in hundreds of a second.
- ITU Capitalisation indicates prominence.
- == an equal sign indicates "latching"; there is no interval between the end of a prior turn and the start of a next turn.
- ??? indicates unintelligible utterances.
- <f> placed below a given syllable or sequence of syllables indicates that they are produced louder than the neighbouring syllables.
- <p> placed below a given syllable or sequence of syllables indicates that it is produced softer than the neighbouring syllables.
- <CRES> placed below a given sequence of syllables indicates that it is produced with increasing loudness.
- <DEC> placed below a given sequence of syllables indicates that it is produced with decreasing loudness.

- <|> placed below a given sequence of syllables indicates that it is pronounced more slowly than the surrounding syllables.
- <a|> placed below a given sequence of syllables indicates that it is pronounced more quickly than the surrounding syllables.
- <H> placed below a given syllable or sequence of syllables indicates that it is uttered with a higher pitch than the surrounding syllables.
- <L> placed below a given syllable or sequence of syllables indicates that it is uttered with a lower pitch than the surrounding syllables.
- % placed at the end of a given segment of speech on the computer print-outs indicates speech unit boundary.