

CHAPTER 4

RESEARCH METHODS

4.1 Introduction

This chapter provides a description of the data used and the method of data collection. It also explains how the data were transcribed, coded and quantitatively and qualitatively analyzed, whilst providing operational definitions of key terms in the study.

4.2 Source of Data

The main consideration for the data in this study was to obtain normal (non-pathological) spontaneous speech produced in a natural, as opposed to experimental, context. Another consideration was to obtain data from a similar speaking context to avoid major situational or contextual differences.

In view of these considerations, data was obtained by recording a particular radio programme over a period of three weeks. The programme is a breakfast show (henceforth, *The Breakfast Show*) that is aired every morning on a local radio station, and has callers phoning in to talk to the two presenters of the programme. This method of data collection is similar to the method used by Blackmer and Mitton (1991). The reason for doing this was that it provided naturally occurring data, which was not affected by studio or experimental conditions. Secondly, since the callers were phoning in to discuss topics suggested by the presenters at that particular moment, it could be

assured that their speech was not prepared prior to their calls. Thirdly, recording callers to the same radio programme over a period of time meant that the data came from the same speaking situation, even if the topics of conversation were not always the same.

The use of telephone data is not new, and has been used in many sub-areas of linguistic analysis, such as Fries' (1952) use of telephone data; the use of the Switchboard and CALLHOME corpus. The use of such data can provide a large sample of natural, spontaneous speech, and is logistically less tedious than tape-recording samples of naturally occurring face-to-face interaction. However, because of the absence of visual cues in telephone conversations, there are bound to be some differences between the features of such conversations and face-to-face interactions. Among the differences that have been found are that the telephone conversations tend to have less pauses and interruptions than face-to-face interactions (Rutter & Stephenson, 1977; Rutter, 1987).

In this study, the speech of 67 callers was used although there were many other callers to the programme during the data collection period. However, their speech was not used in this study for two main reasons. The first was due to audibility. This was especially because many of the callers were using mobile phones, and therefore, there tended to be quite a lot of noise interference. The second reason was that their speech was too short to be used for the purposes of this study. For example, they may have called in to answer a question as part of a competition. There were also studio interviews with local personalities. Since this constituted a different speaking context, these conversations were also not used in this study. Further, since the presenters of the radio programme might have prepared some of their speech, only the speech of the callers was analyzed.

All the 67 callers were perceived as being speakers of Malaysian English. This could be ascertained from the phonological, lexical and syntactic features present in their speech, which were characteristic of Malaysian English, particularly the informal spoken variety. As a transported and transplanted (Moag, 1992) variety, the English used by Malaysians is associated with particular linguistic features which differentiate it from other varieties of English. These include the use of localized vocabulary, distinct phonological features and intonation patterns, as well as differences in syntactic structures and pragmatic features (Anthonysamy, 1997; Baskaran 1984, 1987; Ooi, 1997; Platt & Weber, 1980; Platt, Weber & Ho, 1984; Wong, 1983). Nevertheless, the extent to which the various linguistic features are used differs according to the sub-variety of Malaysian English being used. In other words, like, British or American English, Malaysian English also comprises all the Englishes used by Malaysians from different social and ethnic backgrounds in a variety of contexts (Gaudart, 1997). Thus, factors such as, whether it is a users' first, second or other language; the user's level of proficiency; and the context of use (e.g. audience, topic, purpose, intention, situation, place, roles, degree of formality, etc), can determine the sub-variety of Malaysian English being used.

In view of these different sub-varieties, there is perceived to be a "continuum" of Malaysian English, comprising, according to Baskaran (1994: 27), at least three distinguishable sub-varieties: the acrolectal, mesolectal and basilectal varieties. Placing ME on a continuum of sub-lects suggests that there is no neat division between the three sub-varieties, and indeed users can shift up and down the continuum depending on the types of sub-lects they have in their linguistic repertoire (Morais, 2001). There also need not necessarily be a direct relationship between the type of pronunciation, vocabulary and syntax used in a particular sub-variety. For instance, the English used in more

formal, written contexts displays localized features. An example of this is the use of the word *handphone* instead of *mobile* or *cell phone* in newspaper articles and public notices.

Malaysian English can also be spoken with different degrees of ethnic and geographical accents (Nair-Venugopal, 1997), regardless of whether the variety being spoken is more acrolectal or mesolectal. In other words, the pronunciation and prosody used can range in terms of how ethnically marked (+marked) or unmarked (-marked) it is, where the more marked the accent is, the more prominent the features of that particular ethnic accent is perceived to be (Gill, 1999: 220). Figure 4.1 adapted from Baskaran (1994) and Gill, (1999) shows the continuum of ME, identifying the main linguistic characteristics of the three major sociolects.

| | Acrolect (Standard ME) | Mesolect (Colloquial) | Basilect (Broken) |
|----------------------|---|--|---|
| Syntax | + standard | - standard | extremely simplified structures |
| Lexis | + localized lexical items accepted in formal and informal use | + localized lexical items, including those not used in more formal contexts | Pidgin-like |
| Pronunciation | Can be ± marked ethnically | Usually, but not necessarily + marked ethnic accent and intonation | Usually + marked ethnic accent and intonation |
| Examples | <ul style="list-style-type: none"> • Newspaper reports • Formal letters and documents • Television news • Official speeches | Informal spoken & written communication between colleagues, friends, family members etc. | Used by those with limited proficiency in English |

S = Syntax, L = Lexis, P = Phonology

(from Pillai & Fauziah Kamaruddin, forthcoming)

Figure 4.1
Continuum of Malaysian English

Notwithstanding social and ethnic differences, the common features displayed by the callers included the following features:

- shortening of long vowels, for example in unstressed words like *keep* (C4) and *tall* (C5), which were realized approximately as [kɪp] and [tɔl].
- deletion of final consonants like /t/ and /d/, for example in words like *guest* [C16] and *percent* [C50], which were realized as [ges] and [pə'sen].
- dental fricatives /θ/ and /ð/ being realized as /d/ and /t/, for example in words like *mother* (C44) and *they* (C52), which were realized approximately as [mʌ'də] and [deɪ].
- different pronunciations due to a more syllable-timed stress pattern as perceived in words like *restaurant* (C16 & C56) and *collecting* (C21), which were realized as [res'tɔrən] and [kɔ'lektɪŋ].
- the use of the **is it/ isn't it* question tags, for example, in *ah ikan kembung halia alim is it* (C8)
- semantic shifts of words, for example, *send my children to school* instead of *take* (C12); *opening the pipe* instead of *turning on the tap* (C19); *horn* instead of *honk* (C31); *bang* instead of *knock into* (C32)
- use of the particle *lah* for example, *truelah* (C3); *itlah* [C18]; *answerlah* [C38]; *womenlah* (C49); *breathlah* (C59); *everythinglah* (C64)

- the use of the particle *ah* to indicate a yes-no type of question, often used to confirm information, for example, *baboon ah?* (C37)

4.2.1 Recording

For three weeks, (Monday to Friday), phone-ins to *The Breakfast Show* were recorded using a portable radio-cassette player (*Hitachi 3D80*). During this show, the two presenters encourage listeners to phone in and discuss a certain topic each day. It was unlikely that the callers' speech was edited since the programme was a live transmission. The topics discussed during the recordings were as following:

- i. *women's personalities based on the kind of men they like*
- ii. *the best and worst celebrities in Malaysia*
- iii. *the worst things listeners had done*
- iv. *outcome of a football match between Brazil and Malaysia*
- v. *whether women controlled men*
- vi. *borrowing things from neighbours*

The phone-ins are only a part of the show. The rest of the show comprises songs, conversations between the presenters, studio interviews and quizzes. However, none of these were taken into account in this study.

4.2.2 Transcription

The phone-ins that were considered suitable for use in this study were transcribed using a *Sony Transcriber BM-77*. This was done at the Language Analysis Laboratory of the Department of Speech, University of Newcastle upon Tyne. The conversation between the presenters and the callers were transcribed orthographically. However, only the utterances of the callers were analyzed for the purposes of this study. Most of the transcription conventions used in this study were adapted from Jefferson's transcription system as outlined in Atkinson and Heritage (1984).

Each conversation is identified by a number. The first caller's conversation is identified as Caller 1, the second one as Caller 2 and so forth in the transcripts, and shortened to C1, C2, when they are referred to in this thesis. The speakers are identified as P (presenters) and C (caller). In order to protect the identity of the presenters and callers, the following have been changed in the transcripts:

| | |
|-------------------------|--------------------|
| Name of radio station - | Radio X |
| Name of Programme - | The Breakfast Show |
| Name of Presenters - | P |
| Name of Callers - | C |

Speech in any other language apart from English is indicated in italics. A gloss of Malaysian English expressions used by the speakers was provided, where necessary, in parenthesis in the transcripts. Following Jefferson's system, the transcript does not use normal punctuation conventions, as shown in Figure 4.2, which also illustrates other conventions used in the transcript.

| Notation | Meaning | Example |
|---|---|---|
| < | Beginning of simultaneous utterance | C9 P way to <go dude.> |
| > | End of simultaneous utterance | C <[but] IP <u>but</u> (REP)> we don't have the: er er^ |
| ^ | Interrupted utterance | P door gift. |
| . | Stopping fall tone | C18 C so he ask me to go and get a glass of water. |
| ? | Rising inflection | C24 P what's wrong with you? C I dunno. |
| ! | Animated tone | C51 P oh my God! |
| : followed by time in msec | Silent pause with time in msec | C1 C ... [most m*] IP :364 <u>most men</u> (REP) smell stink. |
| ah er m you know what you call it/that | Filled pauses (including lexical editing terms) | C39 C I told you <i>kasawari</i> is IP er [i*](DEL) IP what you call that ostrich. |
| : following word or segment | Prolongation | C58 C ya I say: IP er Brazil will win seven three. |
| * after segment | Fragment | C33 C [sh*](DEL) IP oh never mind. |
| (()) | Additional information added by transcribers | C12 C ...the <i>bus sekolah</i> ((school bus)) student |
| () | Unintelligible segments – unable to transcribe accurately | C24 C so today I'm returning everything <to you I () the two of you.> |
| [error] | The part of speech that is subsequently altered, repeated or deleted. | C25 C he told me that it's a [ba*] IP <u>it's a smuggled</u> (SUB) one. |
| IP | Interruption Point | |
| <u>repair</u> | Repair following the IP | |
| Type of Repair | | |
| (REP) | repeat | C8 C it's [quite] IP m < <u>quite</u> (REP)>correct. |
| (SUB) | substitution | C50 C [the] IP <u>the</u> (REP) [w*] IP <u>husbands</u> (SUB) is the general |
| (INS) | insertion | C14 C [the] IP <u>in fact</u> <u>the</u> (INS) traffic IP :868 [was] IP er <u>it was</u> (INS) like a bit slowlah. |
| (DEL) | deletion (annotated after the part that is deleted) | C9 C [we don't have](DEL) IP er I mean [that is n*] IP <u>that is a new</u> (INS) concept. |

Figure 4.2

Transcription Conventions

4.3 Coding of Speech Disfluencies

Based on the perceived relationship between speech disfluencies, self-monitoring and self-repair in speech (see Chapter 2 and 3), *speech disfluencies* are defined as *an umbrella term covering all instances where a speaker falters; devices used to interrupt speech; as well as attempts to re-establish fluency* (Lickley, 1998; Postma, Kolk & Povel, 1990; Shriberg, 1999). Based on this definition and the taxonomy of hesitation put forward by Hieke (1981), the disfluencies found in this study were divided into two groups:

- i. ways of pausing or hesitating in speech,
- ii. attempts at repairing disfluent speech.

However, the types of problems encountered by the speaker are not categorized in terms of whether, for example, they are appropriateness, phonological, lexical or syntactic errors. Instead they are seen in relation to the type of subsequent repair made, to ascertain whether the speaker was repeating, deleting, substituting or inserting an item (fragments, a word or words).

As mentioned earlier in this chapter, although the entire conversation between a particular caller and the presenters was recorded and transcribed, only the utterances of the callers were subsequently analyzed. For the purposes of this study, following Du Bois et al (1993), an *utterance* is defined as *a unit of speech corresponding to a statement or a question as indicated by rising or falling intonation respectively*. There can be more than one utterance in one speaking turn. An utterance can also be as short

as a fragment or a word. The perceived intonation was marked using a full stop or period for a stop-fall tone, a question mark for a rising tone and an exclamation mark for an animated tone, as outlined in Figure 4.2. Examples of utterances are shown in the following examples taken from the data.

C1

P that smell good?

C yes.

C19

C nothing. I went to school.

P uhhuh.

C and they were like what did you do to the kerosene?
and I said oh that was that was kerosene?

With reference to the first research question outlined in 1.5, the transcribed utterances of the callers were, firstly, examined for hesitations such as silent pauses, filled pauses and prolongations (Hieke, 1981). Only hesitations that occurred within an utterance (see Table 2.3 for example) were taken into account. Utterance-initial hesitations (see example from Smith & Clark in 2.9.1.2) were ignored as they did not provide enough evidence of self-monitoring and self-repair.

The utterances were then also examined for self-repairs such as repeats, deletions, substitutions and insertions (Lickley, 1998). This was done to discover the patterns of hesitation and self-repairs that emerged from the data. These hesitations and self-repairs were marked in the transcripts and coded as shown in Figure 4.1. The categorization of the data was done at least 3 times over a period of 2 years. This was done to ensure that there was reliability in the way the data was categorized since the writer was the only person doing the analysis for this study.

Instances where only hesitations were present, but no overt evidence of a problem or error and repairs, were considered as *possible-repairs* rather than covert repairs, since it could not be determined if an ‘error’ and its subsequent covert or prearticulatory repair had been made (see 2. 8.2). However, this was with the exception of repeats, which were categorized as self-repairs following Lickley (1998). An example of a possible-repair is given below.

C10

C ... you cannot make it **IP ah** other than *cenderamata*
((souvenir))

In cases like the example above, the hesitation was taken into account, and the interruption point was deemed to begin at the offset of the word preceding the hesitation. In the example above, the interruption point begins at the offset of the word *it*. If the hesitation is a prolonged segment, such as in the example below, then the interruption point is taken to begin at the offset of that segment.

C14

C I think it's **a: IP** free for all.

4.3.1 Hesitations

Hesitations, in this study, refer to pausing devices used by speakers, namely *silent pauses, filled pauses and prolongations*. The following sections describe the different types of hesitations that were coded and analyzed in this study.

4.3.1.1 Silent pauses

Following Lickley (1994, p. 211), *silent pauses* are defined as “periods of apparent silence in the middle of utterance which did not merely coincide with consonantal stop closure or glottal closure”. As shown in Figure 4.1, silent pauses are marked by a colon followed by the duration in milliseconds (see Figure 4.1).

This method of defining silent pauses differs from many other studies, which use cut off points like 100msec, 150msec, 200msec or 250msec to differentiate between phonetic and silent pauses (for example, Arim, Costa & Freitas, 2003; Goldman-Eisler 1968; van Donzel and Koopmans van Beinum, 1996; Zellner, 1994). However, doing this may result in the omission of shorter silent pauses, which have been found as markers of hesitation, in other studies (Blackmer & Mitton, 1991; Kirsner, Dunn, Hird, Parkin & Clark, 2002; Lickley, 1994). Further, as pointed out by Hieke, Kowal and O’Connell (1983), pauses below these thresholds may also constitute non-phonetic pauses (see 2.9.1.1.1).

Parts of the data which contained silent pauses were digitalized at a sampling rate of 12,800 samples per second, using the *Computerized Speech Lab (CSL)* system at the Phonetics Laboratory at the Department of Speech, University of Newcastle upon Tyne. Spectrograms and waveforms of these silent pauses, together with their auditory signals, were then examined using *CSL*. During the later stages of this study, *PRAAT*, a speech analysis and synthesis program available from the internet (<http://www.fon.hum.uva.nl/praat/>), was used, where necessary, to reexamine the silent pauses. These two systems were used because the initial acoustic analysis for this study

was done at the University of Newcastle, while the rest of the study was completed at the University of Malaya.

Similar to Lickley (1994), any absence in the acoustic signal as seen on the spectrogram, and as discerned auditorily were marked, that is, from the beginning to the end of the absence of an acoustic signal. With reference to the working definition of a silent pause adopted in this study, these measurements excluded stop closures. This is similar to the method employed by Lickley (1994). The times for the beginning and the end of the pause as indicated on the spectrogram were noted down and entered into a spreadsheet using *Microsoft Excel*, which was then used to calculate the duration of the silent pause in milliseconds. Figure 4.3 shows the spectrogram and waveform of a silent pause, where the period of silence as indicated by the absence of an acoustic signal.

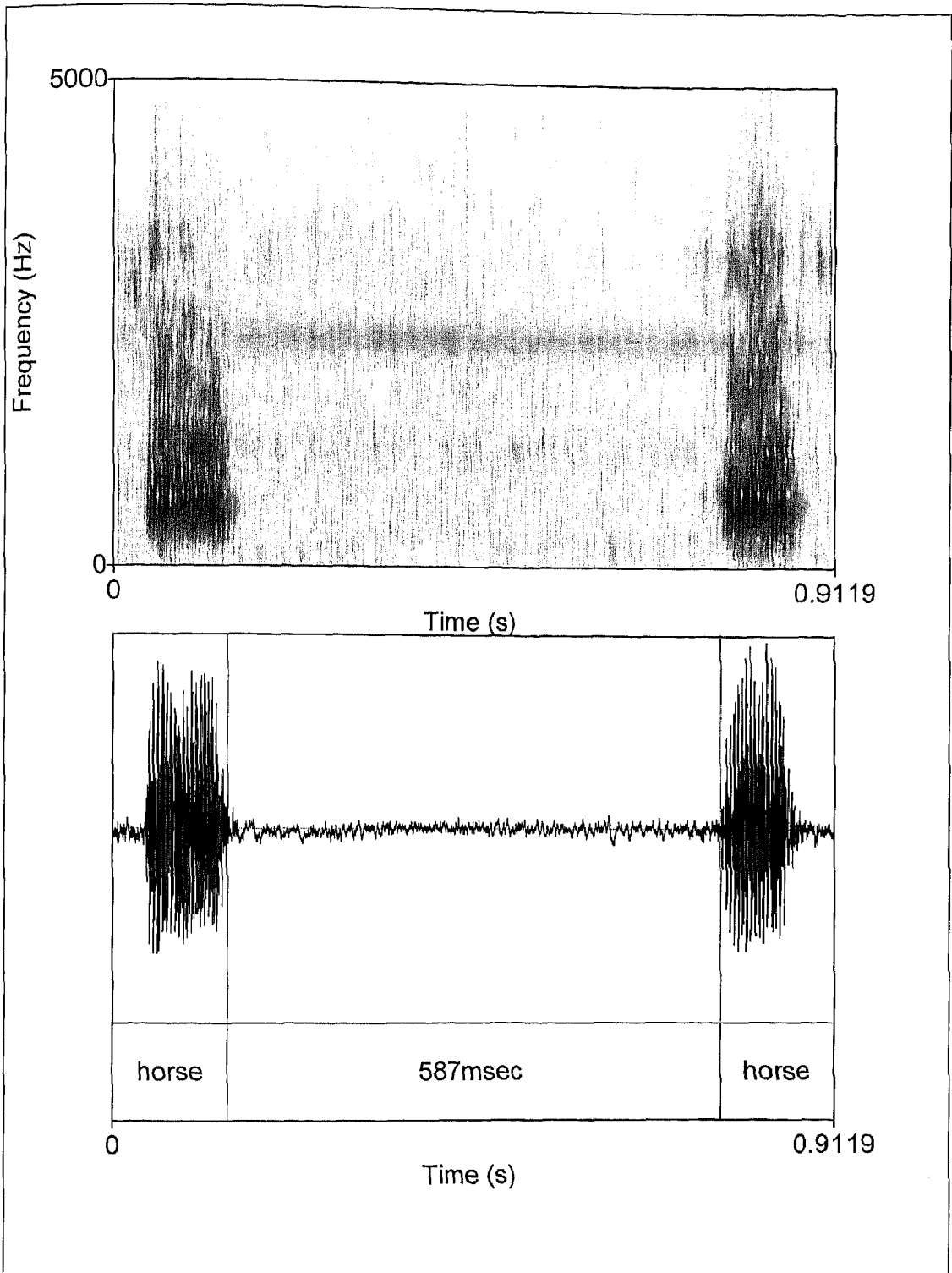


Figure 4.3

Spectrogram and Waveform of a Silent Pause Produced by Caller 11

Following Maclay and Osgood (1959), the location of silent pauses, as well as filled pauses and prolongations, were also examined to determine if they occurred more frequently before lexical words (adjectives, adverbs, nouns, verbs) or before function words (articles, modifiers, prepositions and conjunctions). This was done based on the assumptions that the production of lexical words is more cognitively demanding (see 2.9.1.1.2) than that of function words.

4.3.1.2 Filled pauses

Filled pauses are defined as non-lexicalized and lexicalized fillers that are used in utterances (see 2.9.1.2). Following this, two types of filled pauses were annotated. The first were non-lexicalized vocalizations like ah [ɑ], ahm [ɑm], er [ə] and m [m], (Goldman-Eisler, 1961; Rose, 1998). The second type of filled pauses were *lexical editing terms* (Lickley, 1998), which are considered as lexicalized versions of filled pauses (Abou Haidar, 2001; Rose, 1998), in that they do not contribute to the content of the utterance but actually break its fluency. Examples of lexical editing terms include *you know, what, what you call it/that, what you call this*.

4.3.1.3 Prolongations

Prolongations, in this study, refer to *words within which a sound segment is sustained, usually a word-final vowel*. The duration of a word that was auditorily perceived to be prolonged occurring before an interruption or cut off point was measured by comparing it to the duration of the same word produced in a fluent context by the same speaker

(Shriberg, 1995). This was done by digitalizing the parts of the utterances that contained prolongations, and viewing their respective spectrograms and waveforms, similar to the procedure used to measure silent pauses (4.3.1.1). By using these spectrograms and waveforms and listening to the relevant parts, the duration of both the prolonged and fluent words were measured from the onset to the offset of the word, as shown in Figures 4.4 and 4.5. This was done in all cases except when a stop consonant preceded the prolonged vowel, such as in the word *to*. For this word, the duration of the closure was not measured. Instead measurements were taken from the onset to the offset of the vowel following the release of the stop as shown in Figures 4.6 (the arrows indicate the points at which the duration of the prolongation was measured).

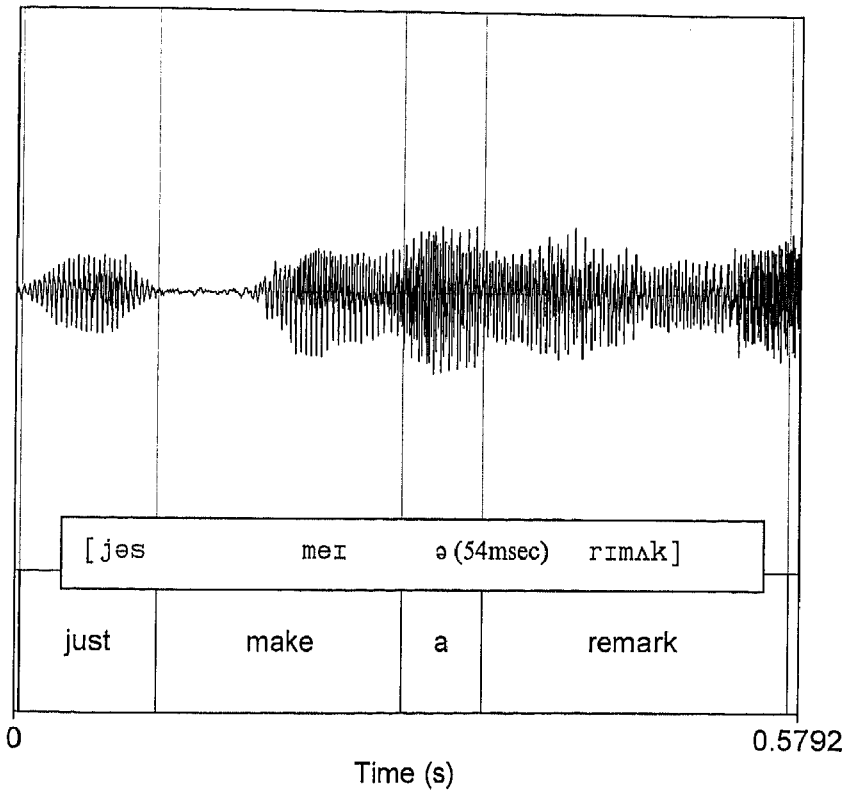
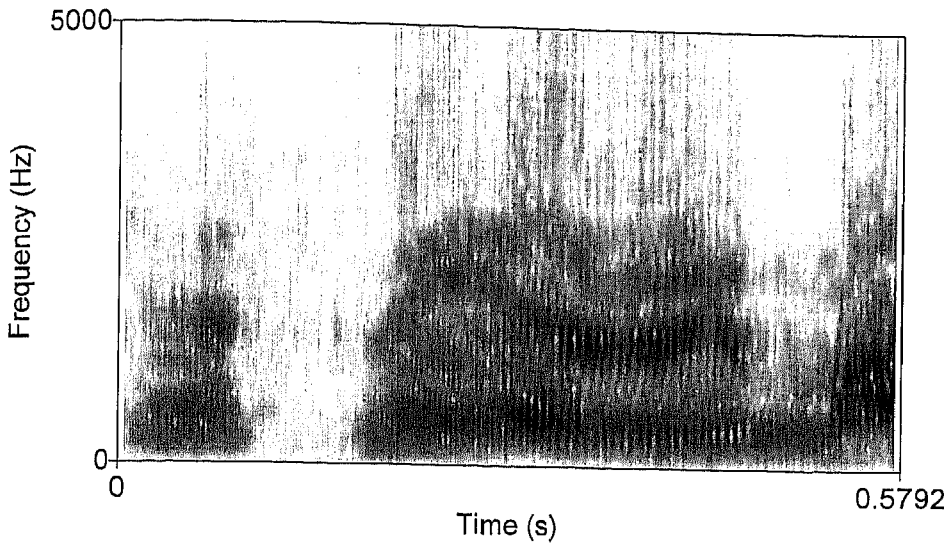


Figure 4.4

Spectrogram and Waveform Showing Fluent *a* Produced by Caller 32

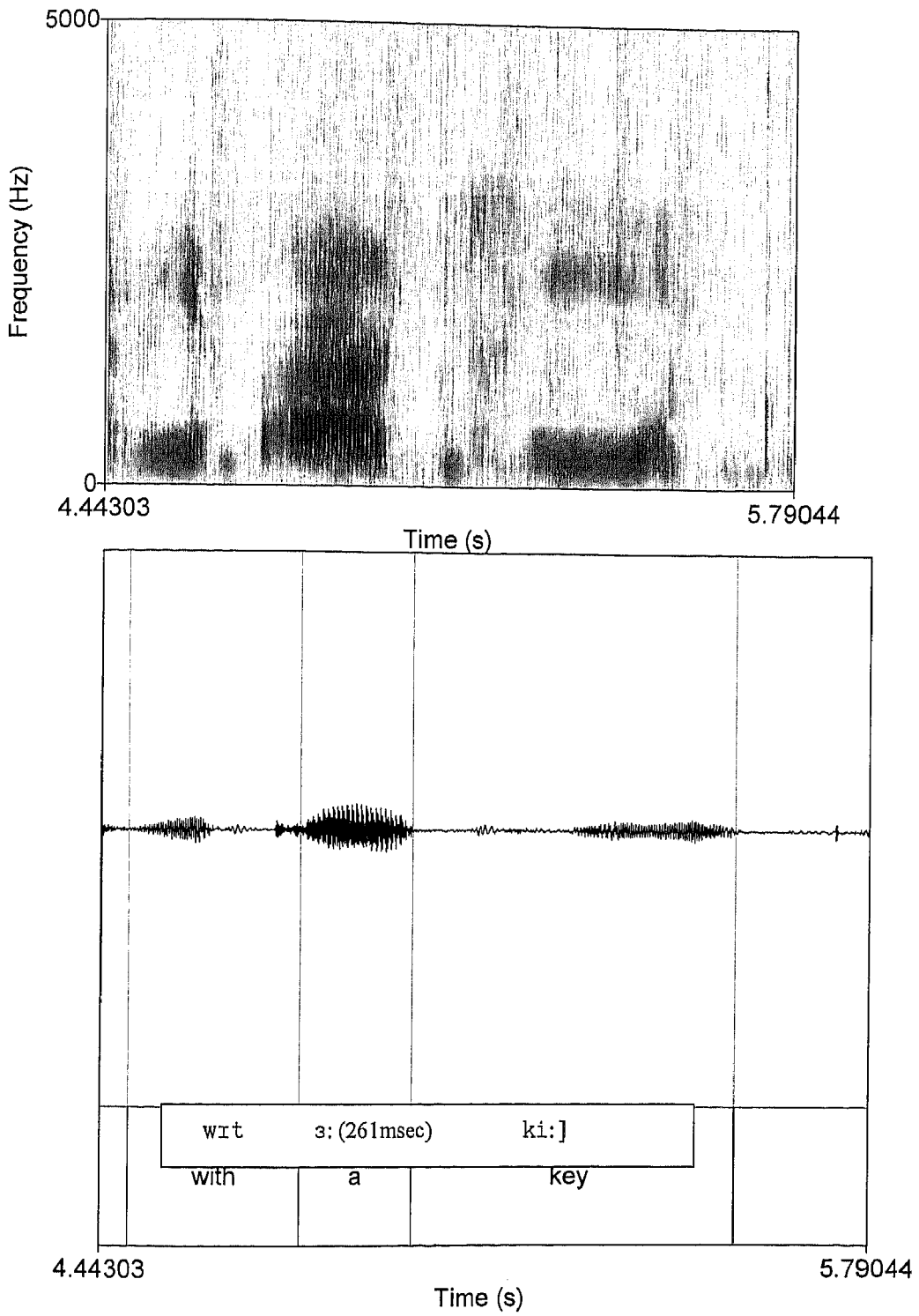


Figure 4.5

Spectrogram and Waveform Showing Prolonged *a* Produced by Caller 32

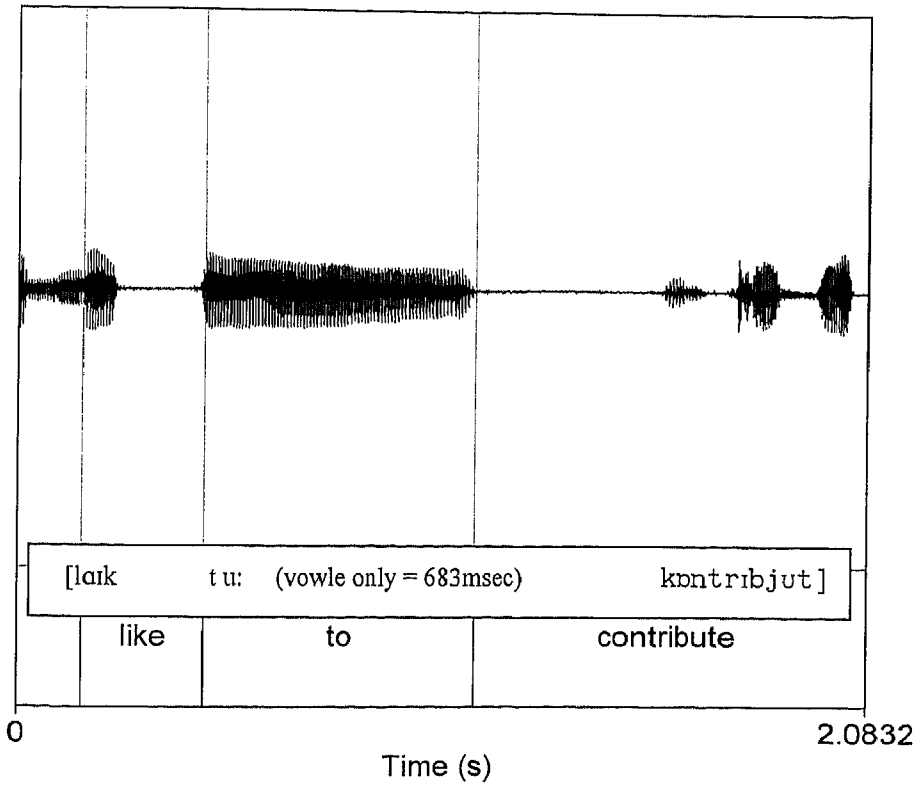
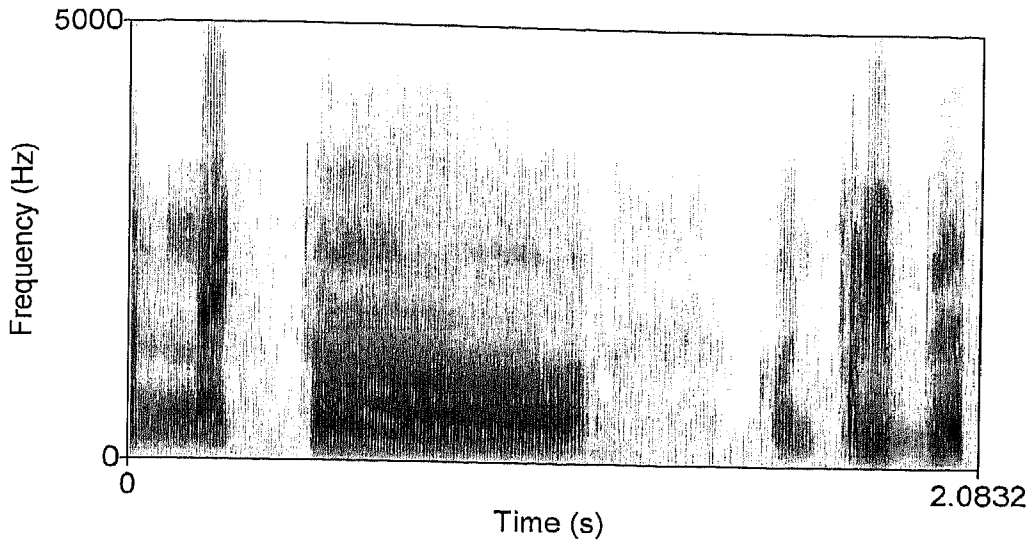


Figure 4.6

Spectrogram and Waveform of Prolonged *to* Produced by Caller 17

It was, however, not possible to compare all the prolonged tokens as some callers did not produce the prolonged word in a fluent context. Thus, the judgement of whether these words were actually prolonged was generally based on how these words are perceived to be pronounced by speakers of Malaysian English. However, as another means of comparing duration, these prolonged words were compared to the same word, which was prolonged by other callers to determine if they were within the same durational range.

4.3.2 Types of Self-Repairs

In this study, *self-repairs* are considered to be *self-initiated repairs*, which occur in response to a problem, occurring within the same unit of utterance in the same speaking turn (Schegloff, Jefferson & Sacks, 1977). The self-repair, therefore, must have a reparandum, which would constitute the error, and its related repair. Following Lickley (1998, 2001) and Shriberg (2001), four types of self-repairs were coded for analysis: repeats, deletions, substitutions and insertions. These were based on the type of corrections made by the speakers. The structure of a self-repair (see 2.8.1), which comprises the error, interruption point, editing phase and the repair, adopted in this study is shown in Figure 4.6 (Levelt, 1983; Lickley, 1998; Shriberg, 1999). However, only the interruption point is indicated if speakers do not appear to overtly make an error and a subsequent repair in their utterance, such as the last example in Figure 4.7.

| Error | Interruption Point | Editing Phase | Repair |
|--------------------------|--------------------|---------------|----------------------------|
| [most m*] | IP | :364 | <u>most men</u> (REP) |
| [I'm calling from:](DEL) | IP | | <u>actually I' driving</u> |
| [in] | IP | ah | <u>near</u> (SUB) |
| [the] | IP | er | <u>one of the</u> (INS) |
| | IP | ah | (Possible Repair) |

Figure 4.7

Structure of a Self-Repair

As illustrated in Figure 4.7, a self-repair is taken to begin from the onset of the 'error'. The beginning of an error for the different types of self-repairs are as follows (see Figure 4.2 for more examples):

- Deletion – from the beginning of the items expunged which are not repeated or changed in the repair
- Insertion – from the beginning of the segment or word before which a word or words are inserted in the repair.
- Repeat – from the beginning of the segment or word that is repeated in the repair.
- Substitution – from the beginning of the segment or word that is replaced in the repair.

In this study, an *error* is defined as the problem that is repaired by way of repetition, deletion, substitution or insertion. As shown in Figure 4.7, the error, therefore, may include words that are not altered or changed but are repeated as part of the repair. An error in this context, also, does not necessarily correspond to accuracy-related errors, such as non-standard pronunciation or syntax. The main criterion is that the error must be related to one of the four forms of self-repairs. In other words, there must be some form of self-repair evident. Following this, the following examples were not considered for analysis in this study as the speakers did not make any self-repairs in relation to the portions that are underlined:

Uncorrected Pronunciation Error

C4

C I was <wondering> **ah IP** what will be my girlfriend be like because I like to keep moustache ((realized as [mʊstʌʃ]))

Uncorrected Grammatical Errors

C12

C <the children always throwing> the rubbish out of the window. ((missing verb to be for the present continuous construction))

C51

C so I don't think so that men are control by the woman. ((unnecessary use of *so* and the article *the*; the past participle form for the passive construction *are controlled* is not used))

The *interruption* or *cut off* point (IP) is taken to be *the point where speech is perceived to be interrupted* (Blackmer & Mitton, 1991; Levelt, 1983, 1989; Shriberg, 1994). The *editing phase*, if present, begins from the offset of the interruption point to the onset of the repair. It can contain within it silent and filled pauses. The interruption point can also be preceded by a prolonged word or a fragment as shown in Figure 4.6. Subsequently, the *repair* begins from the offset of the editing phase, and contains some

form of correction of the error (Levelt, 1989). The repair may be directly related to the error, such as in repeats, substitutions and even insertions, or the beginning of a completely new word or phrase in the case of deletions.

4.3.2.1 Repeats

Repeats are defined as *instances where the speaker repeats a sound, word or words, without making any additions or deletions* (Lickley, 1998, 2001). Following Maclay and Osgood (1959, p. 24), only repeats that were considered as being ‘non-significant semantically’ were annotated and analysed. For example, instances where a speaker was deemed to be using repeats for emphasis were ignored, such as in the following utterance:

C67
C it's very very ((emphasis, not repetition)) <bad.>

As shown in Chapter 2 (see 2.9.2.1), the token preceding the interruption point in a repeat can consist of

- a fragment, that is a *partial repeat*,
- one word or a series of words, that is a *complete repeat* (Tseng, 2003).

This is illustrated in the following examples taken from the data:

i. Fragment or Partial Repeat

C16
R1
 ... met with this nasty [g*] IP
R2
guest(REP).

ii. Complete One-Word Repeat
C9

er the translation is IP er kocak ((shake)) [for] R1
R2
IP for(REP) shake.

iii. Complete More-Than-One-Word Repeat
C53

R1 R2
<[we love women]> IP we love women(REP).

The examples also show that the token preceding the interruption point is referred to as R1, while the one following the interruption point is labelled as R2 (see 2.9.2.1). As explained in Chapter 2, apart from *simple repeats*, where R1 is repeated once, like the ones in the above examples, there could also be *multiple repeats* (Henry & Pallaud, 2003, p.77), where the same token is repeated more than once consecutively. In such cases, the tokens will be labelled as R1, R2, R3 and so forth depending on the number of repeats. This can be seen in the following example, where the second set of repeats are repeated three times (R1 to R3):

C20

well you can improve but not necessary take [the] IP R1
R2 /R1 R2 R3 R4
:305 [the(REP) the the] IP the(REP) fun out of things
you know.

This example also shows that in multiple repeats, the last word is considered as the repair unless there is a fragment or a hesitation between the repeated tokens as in the first set of repeats in the example above. If there are further repetitions of the same

word, then the same criterion is applied. This is further illustrated in the following examples from the data.

C10

 R1 R2 R3
...I check [in in] IP er :376 er in(REP) IP :337 Arabic
dictionary.

C54

R1 R2/R1 R2 R1
[i*] IP [its(REP)] IP its(REP) just that she has [s*]
R2 R1 R2
something(REP)) [you know] IP :56 you know(REP) IP :502
 R1 R2 R3
when she says something of course [we'll we'll we'll] IP
R4
we'll(REP) talk about it but^

In C10, the third *in* is considered as the repair since it is preceded by hesitation. In contrast, in C54, the R2 in the first set of repeats is assumed to be the repetition of the fragment *i**, since speech is halted at this fragment. The second R2 is then considered as the repetition of the first one, which is therefore, both a repair (of *i**) and an 'error' (repaired by the second *its*).

Using the same procedure done to measure silent pauses and prolongations, the parts of the utterances containing repeats were digitalized, and their durations were measured on the spectrograms (from the onset to the offset of the fragment/word/s), with the aid of auditory analysis. However, in order to determine if the repeats were 'retrospective' or 'stalling' repeats based on some of the characteristics proposed by Plauché and Shriberg (1999) in their sub-classification of disfluent repeats (see 2.9.2.1.1), only durations of

simple (Henry & Pallaud, 2003) one-word repeats were analyzed by comparing the durations of the first (R1) and second tokens of such simple complete one-word repeats (R2). Based on the durations obtained and the classification of repeats by Plauché and Shriberg (1999), repeats with longer R1 durations were considered as *retrospective repeats*, while those with longer R2 durations were regarded as instances of *prospective* or *stalling repeats* (see Figure 2.10).

4.3.2.2 Deletions

Following Lickley (1998, 2001), *deletions* are defined as *instances where a fragment, word or words are deleted following the interruption point* (see 2.9.2.4). Fragments and words can be deleted at utterance-initial or within-utterance position. The items deleted must not reappear in the repair as this might constitute a repeat, or insertion (Shriberg, 1994). It should also be determined that the deleted segment or word has no corresponding syntactic or lexical or phonological relationship to a word in the repair as this could be an instance of a substitution. This is particularly in the case of fragments. In such cases, a combination of auditory, acoustic and contextual analysis was used to determine the category of the repair. It is, however, acknowledged that even with the use of such analysis, we can never be a hundred percent certain what word the fragment might have been. Thus, as mentioned previously, the data was re-categorized, albeit by the same person, at different times to cross-check if the same categories were being used. The following examples show the difference between a deletion, repeat, insertion and substitution:

DELETION

C9

C ... [**I want t***](DEL) IP there's ...

(n.b. the entire portion preceding the interruption point is taken to be deleted since there is no indication of the deleted portion being repeated or substituted following the interruption point.)

REPEAT

C16

C met with this nasty [**g***] IP **guest**(REP)

(n.b. the fragment *g** is presumed to be the onset of the word *guest* and is therefore categorized as a repeat.)

INSERTION

C19

C ... both my [**si***] IP elder sisters

(n.b. the fragment *si** is presumed to be the onset of the word *sisters* which appears in the repair following the insertion of the word *elder*.)

SUBSTITUTION

C25

C ...that it's a [**ba***] IP **smuggled**(SUB) one.

(n.b. the fragment *ba** is assumed to be the onset of the word *banned* which is substituted with the word *smuggled*.)

4.3.2.3 Substitutions

Following Lickley (1998, 2001), *substitutions* are defined as *instances where a fragment or word is substituted in the repair* (see 2.9.2.3). Following Bear, Dowding, Shriberg & Price (1993), a substitution is deemed to have taken place if there is an observable semantic or syntactic relationship between the corresponding words before and after the interruption point. However, in the case of fragments, such as in the the example above (C25) we cannot be certain of the fragment in the word we think it is based on auditory and acoustic examination. However, there were only three cases of fragments that were considered to be substituted in the repair (see 6.2.3). Substitutions were subsequently analyzed to see if there was any relationship between the word substituted and its replacement as has been found in other studies (see 2.9.2.3).

4.3.2.4 Insertions

Based on Lickley's (1998, 2001) definition, *insertions* are defined as *instances where the repair is a repeat of a segment, word or words in the reparanda, with the addition of a word or words inserted before one of these words* (see 2.9.2.2). Following Allwood, Nivre and Ahlsén (1990), Shriberg (1994) and Lickley (1998) (see 2.9.2.2), the inserted item could be placed *before* the repeated items in the repair or *within* them. However, no other item in the repair should be altered in any other way as shown in the examples below. The type of words that were inserted were then analyzed to determine why such insertions were made.

C56

C no but [I'm] IP what I'm(INS) saying is...

4.4 Intervals in Self-Repairs

In this study, *self-monitoring* refers to *the process of prearticulatorily or postarticulatorily attending to one's own speech* (Levelt, 1983; 1989), while *self-repair* refers to *overt error correction made by a speaker* (Levelt, 1983, 1989; Postma, 2000).

In relation to the process of *self-monitoring*, *error-detection* refers to the detection of one's own speech problem, such as the errors presented in Table 2.1, either prearticulatorily or postarticulatorily. Within the same context, *repair-planning* in this study refers to the cognitive process of assembling the repair to a prearticulatorily or postarticulatorily detected error.

As explained earlier in this chapter, for repeats, the items preceding the interruption point considered as the 'error', and the one following the interruption point or the editing phase was considered as the 'repair'. Utterances which only have within-utterance hesitation were considered as *possible-repairs*, since there was no direct evidence to suggest that a repair has been made, as reiterated by Hartsuiker and Kolk (2001) and Levelt (1983).

To understand the mechanisms involved in self-monitoring and self-repair, three intervals were measured with reference to the structure of a self-repair as shown in Figure 4.7. These three intervals as shown in Figure 4.8 were:

- i. Error-to-cut off (between 1 and 2)
- ii. Cut off-to repair (between 2 and 3)
- iii. Error-to-repair (between 1 and 3)

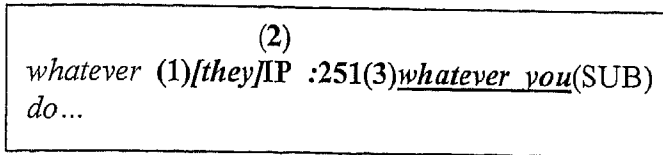


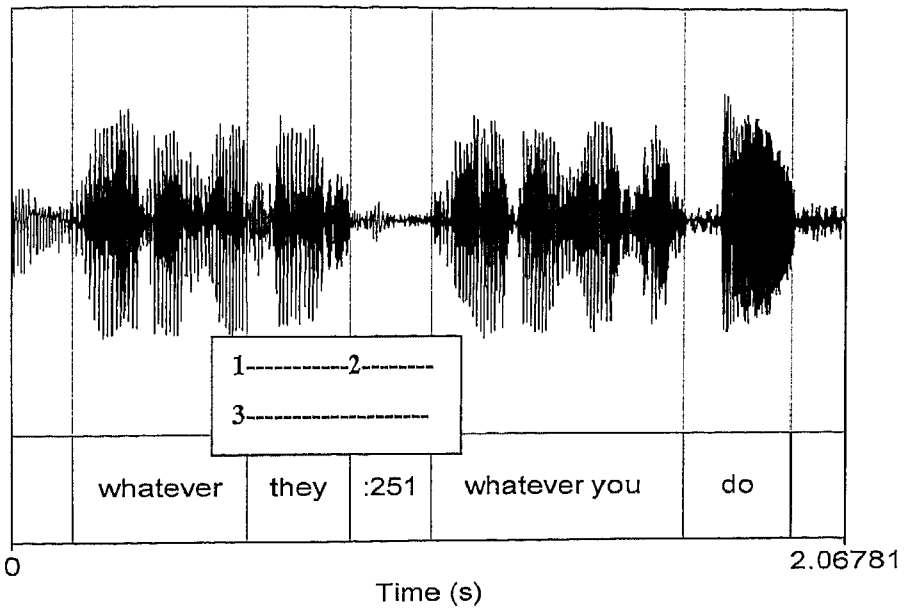
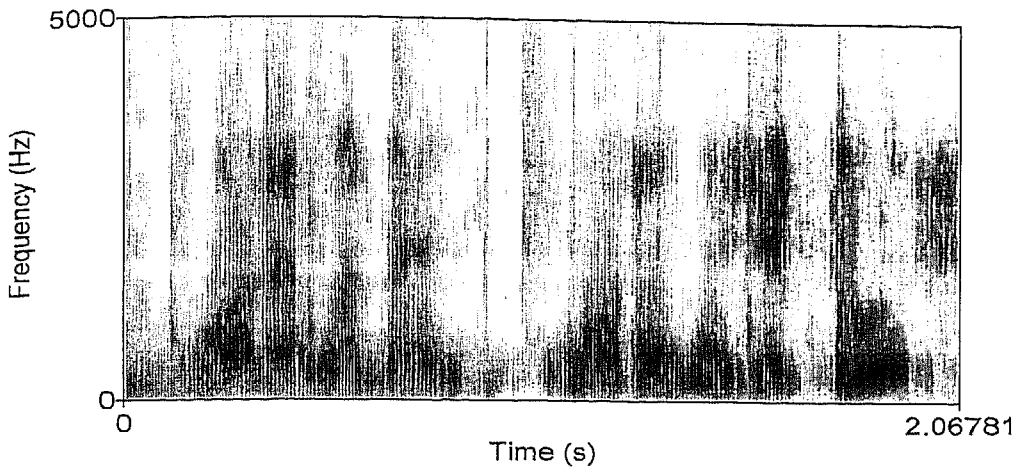
Figure 4.8

Intervals In Self-Repairs

However, for possible-repairs only the cut off-to-repair interval was measured since there was no visible error in the utterance. Following other studies in this area, such as Blackmer and Mitton (1991), Hartsuiker and Kolk (2001) and Oomen and Postma (2001) and van Hest (1996), the term *cut off* instead of interruption point will be used in the context of these intervals (see Chapter 2).

Similar to the method used to determine and measure silent pauses, prolongations and repeats, these three intervals were measured based on spectrograms and waveforms of the repairs, and auditory examination. As shown in Figure 4.9, measurements were taken from the onset to the offset of the error for error-to-cut off intervals (see 1 in Figure 4.9). The repair was considered to begin from the offset of the editing phase, or in the absence of the editing phase, it was taken to begin at the offset of the error. Thus, cut off-to-repair intervals were measured from the onset of the editing phase or offset of the error (or prolongation, if there was one) to the onset of the repair (see 2 in Figure 4.9). Similarly, error-to-repair intervals were measured from the onset of the error to the onset of the repair (see 3 in Figure 4.9). All the measurements were entered into a

spreadsheet and calculated using *Microsoft Excel* and *SPSS*. The descriptive statistics of these intervals were calculated in order to enable comparisons with findings in other studies such as Blackmer and Mitton (1991), Hartsuiker and Kolk (2001) and Oomen and Postma (2001) and van Hest (1996).



(1 = error-to-cut off; 2 = cut off-to-repair; 3 = error-to-repair)

Figure 4.9

Measurements for Self-Repair Intervals

4.5 Statistical Analysis

Measurements of durations of silent pauses, prolongations and the three intervals in self-repair were first entered and calculated using *Microsoft Excel*. These durations were then statistically analyzed using *SPSS*. Descriptive statistics such as the means, medians, modes, standard deviations and skewedness of the durations were obtained. Where, appropriate, histograms and box plots were generated to examine and compare the distributions of these durations in different conditions.

Statistical tests were carried out to compare significant differences ($p < 0.05$) between sets of data only if there were a sufficient number of data. In most cases, non-parametric tests were used because of the following reasons:

- data were not normally distributed
- non-homogeneity of variance between the sets of data
- small sample sizes

The tests used were the

- *Independent Samples T-Test* (to test whether there was any significant difference in the mean scores of two 2 sets of data)
- *Kruskal-Wallis H Test* (to test whether there was any significant difference between the durations of the intervals among the self-repairs)

- *Mann-Whitney U Test* (to test whether there was any significant difference between silent pauses in possible and self-repairs, after fragments and non-fragments in self-repairs, and before lexical and function words in possible-repairs)
- *Wilcoxon's Matched-Pairs Signed-Ranks Test* (to test whether there was any significant difference between the duration of R1 and R2 in repeats)

4.6 Definition of Terms

Although the working definitions of the main concepts in this study have been provided throughout this chapter, the following is a summary of those terms for ease of reference.

- **Cut off or Interruption Point**

Point where speech is perceived to be interrupted. This point may be preceded by a prolongation and/or followed by silent or filled pause.

- **Deletions**

Instances where a fragment, word or words are deleted in the repair.

- **Editing phase**

Begins from the offset of the interruption point to the onset of the repair. It contains within it silent and/or filled pauses, or it can be empty.

- **Possible-Repair**

The production of a prolongation and/or silent pause and/or filled pause within an utterance without any evidence of an overt error and its subsequent repair.

- **Prolongations**

The sustaining of a segment (usually vowel) in a word occurring before an interruption or cut off point.

- **Repair-Planning**

The process of conceptualizing and encoding a repair to an error that has been detected during the process in inner or overt process of monitoring one's own speech.

- **Repeats**

Instances where the speaker repeats a word or words, without making any addition or deletion in the repair. The first token can be a fragment or a complete word. A word or words can also be repeated more than once.

- **Self-monitoring**

The process of checking one's internal and external speech (Levelt, 1983; 1989).

- **Self-repairs**

Self-initiated repairs, which occur in response to a problem, occurring within the same unit of speech in the same speaking turn (Schegloff, Jefferson & Sacks, 1977), and include repeats, deletions, substitutions and insertions (Lickley, 1998, 2001; Shriberg, 2001).

- **Silent pauses**

Non-phonetic and non-grammatical silence mid-utterance, causing a break in fluency determined by measuring the absence of an acoustic signal.

- **Speech Disfluencies**

An umbrella term covering all devices used to interrupt speech as well as attempts to re-establish fluency. Includes hesitation devices like silent pauses, filled pauses and prolongations and self-repairs like deletions, insertions, repeats and substitutions.

- **Substitutions**

Instances where a fragment or word is replaced in the repair and there is perceived to be a phonological, syntactic or lexical relationship between the two.

- **Utterance**

A unit of speech corresponding to a statement or a question as indicated by rising or falling intonation respectively (Du Bois et al, 1993). It can consist of a fragment, one or more words. There can be more than one utterance in a speaking turn.

4.7 Chapter Summary

This chapter explained how this study was carried out, outlining how the data was collected, transcribed and analyzed based on the theoretical understanding of speech disfluencies and the structure of self-repairs discussed in Chapter 2 and 3. The next three chapters will present and discuss the analysis of the data.