CHAPTER FIVE

ISSUES IN SEGMENTATION

5.1 Introduction

The standard approach to the description of intonation especially in the British tradition of intonation is to establish a unit of phonological organisation within which the nucleus or focus can be defined. The assignment of tonal features, in turn, depends on the necessity of having appropriate information points pre-established. The recognition of this central unit in the study of intonation is succinctly described as follows:

"All analyses of intonation postulate a unit which is central in the sense that it provides the framework within which intonational features can be described."

(Scuffil, 1982:34)

They all share a theoretical orientation to define and characterise some units of intonational description, the neutral and unmarked case coinciding with a clause. The intimate relationship between prosody and segmentation of speech is expressed in the words of Gardiner (1977:4) who postulates that intonation segments utterances into
"phrases, signaling to what extent the phrases are related to one another and element within the phrase is the centre of attention".

The fact that it is impossible to utter an extended stretch of speech without some kind of break, and that it is impossible for the hearer to interpret what is said unless what is perceived is chunked into manageable units makes segmentation into divisible units obligatory. Nevertheless, the decision as to how the verbal content of his discourse should be segmented is optional in the sense that it lies with the speaker. The belief that segmentation of discourse is prosodically identifiable and that often (but not always) segmentation is based on speaker decision in pursuit of a purpose provides the impetus for the investigation into the relationship between segmentation and the role that the segmented chunks play in discourse development.

5.2 Theoretical Implications

A question which has to be faced at the beginning of an investigation of this nature is how do speakers segment their discourse. Like prints where commas and fullstops chunk a written text, spoken discourse is also made up of segmented chunks whose manifestations as segmented chunks are brought about by the presence of certain prosodic cues.
The use of different terms to refer to varying forms of segmented speech: tone-groups, tone-units, breath groups, intonation units, intonational phrase, phonemic clauses and pause-defined units suggests that there is some kind of consensus among linguists that a speaker segments his stretch of speech into divisible units. Despite the differences among the aforementioned theoretical constructs, Ladd (1986:311) postulates that they share common characteristics among which are as follows:

(1) they are the largest phonological chunk into which speech is segmented, extending from one phonetically defined boundary to the following,

(2) they have specifiable intonational structure, including in most versions of the theory - a single most prominent syllable (tonic/nucleus), marked by rhythmic and pitch prominence

(3) they are phonological units which are nevertheless assumed, ideally, to match up with elements of syntactic or discourse-level structure.
According to Tench (1990:22), the use of these varying terms "goes to show that we segment events into units and we use this segmentation as an operation by which we divide up ideational content of our communication into manageable parts".

5.3 Problems with Tone Group Boundaries

The ease with which tone group boundaries can be unambiguously determined appear to vary considerably among analysts. According to Crystal (1969) the tone group (which he calls tone unit) is identifiable by its internal characteristics, such that minimally it contains an obligatory nucleus marked by a major pitch movement, or it may consist of three other optional elements - the Prehead (any initial unaccented syllables), Head (any accented syllables preceding the nucleus) and Tail (any unaccented syllables following the nucleus). The boundaries of the tone-group, in turn, are marked by the presence of audible phonetic cues (see Section 5.4 of this chapter).

Like Crystal, Pierrehumbert (1980) identifies the unit of intonation which he calls intonational phrase by its internal characteristics. An intonational phrase is bounded on each side by boundary tones (initial, which is optional and final, which is obligatory), at least one pitch accent which is obligatory and a phrase accent. It is assumed that, in general, an
intonation phrase boundary is marked by the presence of a non-hesitation pause or where "a pause could be felicitously inserted without perturbing the pitch contour" (p. 7). In natural speech the lengthening of the last syllable in the phrase may replace pause as the phonetic cue for segmentation. Pierrehumbert's finite state grammar which "generates the set of well-formed tonal sequences for an intonation phrase" (p. 13) is represented in the following transition network:

![Transition Network Diagram]

In theory, tone groups are supposed to be set off by audible prosodic cues at the actual boundary. They are also supposed to have well-defined intonational structure which correspond with units of grammatical constituent, though not in a predictable way. It appears that
Crystal regards this assignment of tone group boundaries as a positive virtue "any process of intonation analysis will take simultaneously account of both boundary cues and internal structure, and any comprehensive definition of the tone-unit must also have recourse to complementary cues" (1969:205). The problem with this kind of approach is that it is simply assumed that such boundaries identify the domains within which intonational features are described.

In practice, however, tone group boundaries are sometimes not set off by audible phonetic cues and that segments which are demarcated by audible breaks may not contain a nucleus (see Knowles 1991). Crystal (1969) himself admits that he sometimes take recourses to syntactical or semantic criteria to place boundaries. While Crystal (1969), and Halliday (1967) do not seem to face much difficulty in segmenting continuous discourse into tone units, Brown et al (1980:29) occasionally encounter problems.

What seems to make segmentation difficult is the presence of hesitation phenomena brought about by planning and production difficulties. The very high proportion of performance errors in Brown et al's (1980) data of spontaneous informal conversations, e.g. false starts, hesitation, slips of the tongue, overlapping and incomplete utterances contribute to making segmentation problematic. The difficulty that the
speaker faces in producing what he wants to produce may force him to pause or lengthen a syllable at an inappropriate place, thereby disrupting the prosodic flow of the utterance. This break in the prosodic flow causes a stream of speech to be realised as consisting of fragmented chunks whose boundaries may not coincide with syntactic boundaries nor intonational boundaries. For example, elements of close-knit syntactic constituents may be separated from each other by a prosodic break. A speaker may pause before reaching the nucleus resulting in segments which have no nucleus. Although intonationally insignificant, these planning units are important in the sense that they fulfill the speaker's interactional purpose, i.e. gaining him time to prepare his subsequent contribution.

5.4 Pause-defined Units

To ease segmentation, Brown et al (1986) propose dividing stretches of speech according into pause defined units, chunks of speech marked by pauses of more than 0.6 to 0.8 seconds (1980:56). Although highly reliable scientifically, Couper-kuhlen considers pause-defined units unsuitable because many pauses are "performance-related" and as such "do not necessarily reflect the (competence-based) intonational structure of speech" (1985:75).
Brazil et al (1980), on the other hand, skirts the issue of tone group boundaries by adopting an approach in which prominence and tonic mark the centre of a unit whose boundaries are unclear. Without denying the importance of tone group as a basic unit of intonation, Brazil et al (1980:45-46) regard identification of tone group boundaries as of limited significance. For Brazil (1975, 1978, 1981) semantic distinctions come first in intonation analysis.

Brazil et al (1980), however, did not face as much difficulty as Brown et al (1980) in segmenting their data of spoken discourse into tone groups. A plausible explanation for this is the kind of data chosen to form the basis of analysis, i.e. classroom talk. Being more controlled in its context and content, spontaneous speech produced in a classroom situation is often fluent and less marked by hesitation phenomena. It is partly for this reason that broadcast interview has been chosen to form the corpus of data for analysis. Couper-Kuhlen concludes that, despite

"admitted difficulty, it seems wiser to persist in the search of tone groups and tone-group boundaries based on phonological/ phonetic criteria, making allowance for cases of indeterminacy".

(p. 76)

What Couper-Kuhlen says seems reasonable because in any kind of study that deals with something as subjective and unpredictable as spoken discourse, one cannot avoid uncertainty and ambiguity. Since speech is
performance that takes place in real time, one has to accept that "errors of performance" will be present during that act of performance. What is required is a theoretical framework that will, to a certain extent, ease segmentation.

5.5 A Proposed Theoretical Framework

Noting the difficulty involved in identifying tone groups in stretches of spontaneous speech, this chapter does not intend to divide discourse into separate tone groups. Rather attempts will be made to identify the prosodic cues that contribute to making a stretch of speech hearable as consisting of fragmented chunks. The methodology for identifying the segmented parts thus rests on the identification of audible prosodic features whose presence contributes to making the units they bound hearable as segmented chunks. The question of whether these segmented parts contain a nucleus or not is not at issue here because the aim is not to describe the intonation patterns of Malay. After these speech segments have been identified, the contribution that they make to the development of discourse will be examined. The speech unit may be demarcating a unit of information with an item of information being focused as in a segment containing a nucleus or a unit of planning (for example, a filled pause as a separate segment) or a unit of performance error (for example, a slip of the tongue, a false start, an incomplete lexical item, etc.).
Bolinger (1989) regards segmentation (which he refers to as "the demarcative function of prosody") as "universal" such that "all languages use intonation, rate, and pause to mark divisions" (p. 82). For example in Cayuvava, "a fall and a trailing off after final accent" signals the wish not to continue and "a maintenance of the level of the final accent" indicates incompletion (1989:82). According to him a speaker indicates a break in the discourse by a shift in pitch, in particular a drop in pitch. Likewise Brazil (1975, 1978), Lehiste (1975) and Bolinger (1989) note that the speaker indicates the beginning of a new utterance or topic by a change in prosody i.e. raising the pitch height and/or increasing loudness.

5.5.1 Audible Prosodic Cues

Although most linguists recognise the tone group as the basic intonational unit, only Crystal (1969:204-7) gives the most complete discussion of the phonetic cues which signal the boundaries of tone group. He claims that in most cases its boundaries can be determined by phonetic/phonological cues (1969:206) which are as follows:

1. audible change in pitch level at the boundary depending on the direction of the nucleus,

2. the presence of audible pause, final syllable lengthening or aspiration at the end of tone-group.
Crystal further adds that in the absence of these phonetic cues one can then resort to relying on "grammatical or semantic criteria to place the boundary" (p. 207), but such cases are few.

Couper-Kuhlen (1986) and Cruttenden (1986) list out the following external criteria for boundary identification: pause location (either real or potential), final syllable lengthening, rhythmic discontinuity and the presence of anacrusis. Another boundary signal is the relative tempo of unstressed syllables.

To overcome the problem of segmentation, Ladd (1986) proposes the recognition of two types of intonational phrasing: Major Phrase (MP) and Tone Group (TG). The former is delimited by the presence of overt phonetic cues such as pause, anacrusis, syllable lengthening, pitch change, change of tempo, etc. whose presence are agreed upon by most linguists. Tone group is "merely a structural unit of phonology - the domain within which nucleus is defined", whose existence are identified solely on the basis of tonal structure (i.e. the presence of nucleus). These two types of domains form hierarchical structure and that they are recursive in the sense that a constituent is allowed to dominate a constituent that is higher in rank, parallel to the situation in syntax. In Ladd's opinion this theoretical framework will help overcome problem in cases where no phonetic cues
bound segments whose internal structure satisfies the minimum requirement of an independent tone group.

Pike presents (1962) a detailed description of how the boundaries of phonological units can be identified using rhythmic criteria. According to Pike while prenuclear unstressed syllables are characterised by crescendo loudness and faster tempo due to the relative shortness of the syllable, the post nuclear syllables are marked by decreasing level of loudness (decrescendo), lenis articulation (including devoicing) and relative lengthening of the syllable. O'Connor (1973) adds that one can distinguish prenuclear unstressed syllables from the post nuclear unstressed syllable by relative tempo such that the former is uttered relatively faster than the latter.

At this juncture it is worth drawing together the various observations that linguists have made as regards the phonetic criteria that contribute to the realisation of stretches of speech as separate units of prosody. They are summarised in Table 1 below:
### TABLE 1: External Cues to tone group boundaries

<table>
<thead>
<tr>
<th></th>
<th>Pitch Parameters</th>
<th>Tempo Parameters</th>
<th>Other Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crystal/Crutenden/</td>
<td>change of pitch level/ pitch direction</td>
<td>pause final syllable lengthening</td>
<td>semantic/syntactic criteria</td>
</tr>
<tr>
<td>Coupen-Kuhlen/</td>
<td></td>
<td>anacrusis rhythmic discontinuity</td>
<td></td>
</tr>
<tr>
<td>Ladd</td>
<td>pitch change</td>
<td>anacrusis final syllable lengthening pause</td>
<td>major group identified by the presence of audible prosodic cues</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>minor group identified by the presence of nucleus</td>
</tr>
<tr>
<td>Pierrehumbert</td>
<td>boundary tones</td>
<td>pause final syllable lengthening</td>
<td></td>
</tr>
<tr>
<td>Pike</td>
<td></td>
<td>relative allegro tempo of pre-nuclear unstressed syllables</td>
<td>crescendo loudness of pre-nuclear unstressed syllables</td>
</tr>
<tr>
<td></td>
<td></td>
<td>relative lengthening of post nuclear unstressed syllables</td>
<td>decrescendo loudness of post nuclear unstressed syllables</td>
</tr>
<tr>
<td>O'Connor</td>
<td></td>
<td>prenuclear unstressed syllables uttered relatively faster than post nuclear unstressed syllables</td>
<td>lenis devoicing</td>
</tr>
</tbody>
</table>
5.6 Truncated Segments

Based on the external cues described above, a preliminary analysis of the data reveals the occurrences of the following truncated segments whose audible prosodic cues at the boundaries give them the auditory effect of being cut off or incomplete. This incompleteness is brought about by the fact that these speech units are syntactically, semantically and/or intonationally incomplete. They can be categorised as follows:

(1) incomplete speech units whose boundaries are set off by prosodic cues. Syntactically and semantically they are part of the subsequent chunk of speech from which they are separated.

(2) abandoned speech units whose realisation is the consequence of unfluent speech. Unlike the first, the speaker leaves the segment incomplete and starts fresh. The boundaries are demarcated by audible prosodic cues. They do not cohere syntactically or semantically with the segment following them.

(3) parenthetical speech units which interrupt the prosodic flow of the primary utterance. The boundaries of these "wedged in" segments are set off by prosodic cues.
5.6.1 Incomplete Segments

Extracts 5.1 through 5.3 present a dilemma in intonation analysis. It does not require a very close examination to notice that the highlighted truncated segments of speech (arrowed) in Extracts 5.1, 5.2 and 5.3 below share similar characteristics such that they are marked off from the segments following them by a pause and they are intonationally incomplete in that they lack a nucleus. The speaker breaks off before reaching the nucleus, a point that is illustrated by the absence of the nucleus (noted by the absence of capitalisation).

**Extract 5.1**

1B kita more INDIVIDUALISTIK (0.23)

< (0.76)

< (0.63)

-> dia: (0.42) daripada (0.43)

< (0.55) < (0.35)

< (0.41)

< >L

< >p
AWAL Jlagi TADI: dididik secara berkumpul
< >(1.15)
< >(0.46)
< >(1.2)

B We are more individualistic. From young, they are brought up in groups.

Extract 5.2

2A how is it different
< >(0.93)

-> ataupun ada (0.29) KESAMAAN
< >(0.58)
< >(0.27)

A Is it different or is there any similarity?

Extract 5.3

-> 3B kita (0.36) pun tukar MODEL
< >(0.25) < >(1.2)
< >L < >L
< >P < >F

B we then change model
Although the lack of nucleus contributes to making the highlighted segments intonationally insignificant, they are significant in the sense that the pause which bounds them affects the prosodic flow of the discourse and causes discontinuity. Pauses of this type, are generally regarded as hesitation phenomena whose occurrences are quite common in our data. The relative lengthening of the final syllable further supports the assumption that the speaker is planning what to say next.

Despite the break, the speaker maintains a link with what follows by producing the final syllable with a level pitch which neither falls nor rises, thereby communicating the intention to continue. In Extract 5.1, although the overall pitch of "dia" (they) is low, the speaker signals the wish to continue by keeping the pitch level and lengthening the final syllable. One can easily detect that the speaker is having planning difficulties by his relatively slower speech rate, i.e. 550 msec. in producing a two syllable utterance in comparison with 350 msec. that he takes to utter a four syllable utterance located subsequent to it. The rush to produce "daripada" (from) causes the speaker to pause to replan what he wishes to say. Again he indicates incompleteness by producing the final syllable with a non-low pitch level. Besides the pause which causes "daripada" to be broken off from the subsequent talk, the speaker also distinguishes the latter from the former by a step-up in pitch when producing "awal" (young), the initial word in the subsequent segment.
Likewise in 5.2 and 5.3 "ada" (there is) and "kita" (we) respectively are separated from the subsequent segments by a brief pause, yet are prosodically bound to them by their terminal pitch which stays non-low. This cue is used by the speaker as a floor keeping strategy while planning his utterance.

Gumperz (1992b:235) calls this type of pitch that suggests incompleteness "a holding intonation". In a recent study on the management of talk, Local (1992:275) observes that these pitch characteristics along with loudness are used by speakers to hold turn or project more talk by the current speaker (see Chapter Eight of this thesis).

It is inevitable that in producing spontaneous speech speakers may produce segments whose prosodic boundaries do not coincide with syntactic boundaries. This is so because unlike the latter, phonetic entity is basically some sort of unit of performance. Speakers, for example, pause in the middle of noun phrases, verb phrases and other close knit syntactic constituents, changing pitch height and direction on syllables for the purpose of keeping the floor while planning what to say next. In Extracts 5.4-5.5 below, a close knit noun phrase "dua ketul besi" (2 pieces of iron) (indicated with an arrow) is separated from each other by a pause, defying syntactic cohesion, and in Extract 5.6 a modal, "boleh" is separated from its verbs, "membawa" (bring).
Extract 5.4

4B kalau kita `AMBIK (0.32)
-> a: DUA (0.24) ketul BESI%
   < > CRES

B If we take two pieces of iron

Extract 5.5

5B dan yang `tua baru `DARJAH (0.28)
   < > (1.46)

`EMPAT
   < > (0.34)
   < > H
   < > CRES

B And the eldest is only in standard four.

Extract 5.6

6B Maknanya kalau saya `BOLEH (0.34)
   <F F>

a: `MEMBAWA (0.32) a:
   < > (0.56)
seorang `JEPUN (0.28)

< f >

ke `MALAYSIALAH

<f f >

B Meaning if I can bring a Japanese to Malaysia

In 5.4 "dua" (two) is separated from a quantifying noun by a brief pause. The final syllable is relatively high signaling incompletion and the word is uttered with crescendo loudness. Likewise in 5.5 the last syllable of "darjah" (standard) is uttered with a higher pitch than the preceding one and accompanied by forte loudness.

In 5.6, the speaker's short utterance is segmented into four prosodically marked segments. The separation of the prior segment from the subsequent one signaled by a brief pause is done at a syntactically inappropriate point. The word "boleh" (can) is uttered loud and with a non-low pitch height signalling incompletion. The occurrence of "a:" whose level pitch seems to serve no function other than to maintain the speaker’s turn while he plans his utterance further supports this assumption.

These "planning units" as Cruttenden calls them (1986) are the realisations of planning problems which often occur in natural spoken data and are usually separated from the corrected utterance by a brief pause.
According to Brown et al (1980:68) these "search pauses" which are brief in duration (i.e. between 0.28 and 0.38 seconds) are the results of planning problems and such an assumption is correct because the units whose boundaries they demarcate are incomplete syntax.

5.6.2 Abandoned Segments

Another type of truncated unit identified in our data includes what is referred to as "abandoned segments" which is made up of incomplete syntax. In our data what is observed is that the speaker usually goes back to the beginning of such segments and reproduces them again, this time more fluently, e.g. "mana jum ..." in 7A, "maksud sa ..." in 8A and "bukanan ..." in 9A:

Extract 5.7

7A  `MAAFlah (0.15)
   <  >(0.43)
      gurau sikit'DOKTOR (0.3)
      <  >(1.03)

a-> mana jum ... (0.21)
   <  >(0.39)
b-> ^mana jumpa Jepun tu
< > (0.41)
< L L> (0.59)
<f f>
A Sorry. Just joking doctor. Where did you meet the Japanese?

Extract 5.8

a-> 8A maksud sa (0.40) ...
< > (0.43)
< DEC >
<L L>
^maksud saya:
< > (0.51)
<f f>
sekarang ni kereta, PROTON
< > (0.34)
< > (0.39)
A What I mean nowadays PROTON cars
Extract 5.9

9A  **bukanan (0.21)% ...**

\[
\begin{align*}
\text{bukan nak congki} \cdot \text{PERIBADI} & > (0.36) \\
\text{(0.97)} & > \text{H} \\
\text{f} & \text{f}
\end{align*}
\]

A  Not that I want to ask about your private life.

In Extracts 5.7, 5.8 and 5.9, the speaker produces an incomplete utterance, pauses briefly and proceeds to repair the preceding contribution by repeating it and thereon continues with the utterance. In Extract 5.7, the speaker distinguishes the speech "error", "mana jum ..." ((arrowed a), where...) from the "corrected" segment "mana jumpa Jepun tu" ((arrowed b) where did you meet the Japanese) by uttering the beginning of the repair with a pitch step-up and producing the latter relatively louder than the former. This prosodic marking is also accompanied by a shift in tempo whereby the corrected utterance is uttered at a faster pace than the prior talk, i.e. the speaker takes 410 msec. to utter a seven syllable utterance as compared to 380 msec. to produce a three syllable utterance.
In 5.8, again the change in prosody contributes to the hearing of a stretch of speech as two segmented chunks. The speaker signals the start of a new contribution by a marked increase in loudness and a pitch step-up from a low pitched "sa" to a relatively high-pitched "mak". The brief pause which contributes to the abandoned segment hearable as a segmented portion gives the speaker time to reformulate his utterance.

In Extract 5.9, the contribution which initiates repair "bukan nak" is separated from prior segment by a brief pause and a noticeable shift in prosody. The repair begins with a correction of the mispronounced "bukanan" in combination with a marked increase in pitch height and loudness.

The self corrections give the impression that the speaker is consciously aware that he has not got right what he was trying to produce and he repairs it. When the repair is made, the speaker contextualises it with prosodic cues that indicate its status as a new contribution.

Sometimes, instead of correcting the false start by repeating it, the speaker abandons this incomplete segment and proceeds to make a fresh start with a new utterance. Extracts 5.10 and 5.11 lend support to such occurrences:
Extract 5.10

9A juga: (0.2) ...

< >L
< >p
^DAN lagi SATU
< >H
< f >

A also and one other thing

Extract 5.11

10A bila DISEBUT a:: (0.24)...

bes (0.36)

<L>
<p>
LOGAM
<H H>
<f f>

A When it is uttered metal
Extract 5.12

11A dan ini ada: (0.32)

< L L>

< > ral
< > creaky
< >(0.56)
< > DEC

^boleh saya sebut Doktor Sakinah pagi ini

<al al> 0.29

A And there is ... Can I mention, Doctor Sakinah, this morning?

The cues that contribute to making the utterances in 5.10, 5.11 and 5.12 hearable as being made up of segmented portions are pauses which divide the prior talk from the following one, pitch shift and a change in volume. Being prosodically incomplete segments in the sense of having no nucleus "juga" and "besi" are therefore prosodically part of the new contribution that follow them. In 5.10 the speaker makes a false start in uttering "juga" (also) detects the error abandons it and starts fresh. The correction, "Dan lagi satu" (and one other thing), begins with a step up in pitch accompanied by forte loudness. In 5.11 the speaker changes his mind in the middle of uttering the word "besi" (metal), and after a pause of 360 msec. he replaces it with the word "logam" which is distinguished from "besi" by a relatively higher pitch and volume. Likewise, in 5.12, A
produces "*dan ini ada*" (and there is) changes his mind and reproduces "*boleh saya sebut Doktor Sakinah pagi ini*" (can I mention, Doctor Sakinah, this morning). The abandoned speech unit is set off from the new contribution by a pause of 320 msec., a pitch step up to "*boleh*" (can) and a change in tempo from lento on "*ada*" (there is) to allegro on "*boleh saya*" (can I).

Syntactic and prosodic discontinuity of a stream of speech caused by the presence of incomplete portions within it is an indication of a hesitant speech. For instance in 5.13 below, one can easily see that the speaker is facing some kind of difficulty in formulating his utterance. His discourse is broken up into five segmented chunks out of which three are incomplete and abandoned (arrowed a, b and c respectively). A plausible explanation for this is that the focus of the talk is on matters which are rather personal, i.e. concerning the interviewee's marriage with a foreigner and how he has managed to overcome cultural barriers by marrying her. One can detect the interviewer's uneasiness with this topic by the number of times he has attempted to reformulate his utterance which results in incomplete and abandoned segments and the interviewer's apologetic remark "*ini maaflah doktor ya*" (I must apologise for this doctor).
Extract 5.13

12A bagaimana kita BOLEH (0.34)

< H H >

< CRES >

< >(1.0)

a-> a: dapat menawan hati: (0.28) ...

< >(0.44)

< >(1.4)

b-> tibukanlah kita: (0.31)...

<LL>

<>H

< DEC >

< >(0.4)

< >(1.07)

c-> tselalanuya kita kalau macama (0.42)...

<al al> <l l>

< > (0.31)

< > (1.09)

ini maaflah doktor ya (0.23)

<H H ><L L>

<f f >

< > (1.23)
A How we can ... are able to win the heart. Not that we ... Usually we if
like ... I would like to apologise for this doctor.

In Extract 5.14 below, only after two unsuccessful attempts (i.e. "tapi
dah" (but has) and "yang dok" (which doc)) is the speaker able to phrase
his utterance successfully (indicated by an arrow):

**Extract 5.14**

13A tapi: dah (0.23)% ...

< (0.7)

a-> tyang dok (0.42) ...

< (0.61)

b-> tyang doktor kata TADI

< (0.92)

<HH>

<f f>

A but has ... that doc ... that you said just now.

The speaker's utterance contains two syntactically incomplete
segments, i.e. "tapi dah" and "yang dok", bounded by a brief pause 230
msec. and 420 msec. respectively. The first segment is abandoned and
left syntactically incomplete. The false start in the second attempt at
producing his utterance further suggests that the speaker is having trouble
with a formulation. The speaker initiates repair by uttering the segment again and this restart is distinguished from the preceding talk by a step up in pitch and loudness on "yang" ((that), arrowed b). In this example the repair is also timed at a faster pace compared with the prior segment. The difference in speech rate between the prior and following talk serves to signal and in part to constitute their different status. The boundaries of the two incomplete segments are demarcated by syllable lengthening and a brief pause.

5.6.3 Parenthetical Segments

Parenthetical segments are brought about by the insertion of a secondary utterance within a primary one causing the latter to be divided into segmented chunks. Parenthesis requires clear prosodic chunking so that the embedded portion can be distinguished from the segmented primary portions. A preliminary analysis of the data reveal a number of these disruptive segments whose boundaries are clearly demarcated by audible prosodic cues.

A common example of this type of "wedged in" segment found in our data is one that is introduced for the purpose of ensuring that the turn is kept yet allowing the speaker time to plan what to say next. This "turn-keeping" and "buying-time" strategy is clearly reflected in the following
extracts whereby the speaker uses "wedged in" phrases like "apa ni" (what's this) in 5.15, "apa tu" (what's that) in 5.16, "apa nama" (what is it called) in 5.17 to signal that he is in search of words to complete his prior talk and he still intends to continue.

Extract 5.15

14A jadi: masa saya::

< >(0.55)

< >(1.22)

-> apa ni: (0.32)

< >(0.44)

< >(0.76)

<H> < >L

< >DEC

†attachment TRAINING

<H H>

<F F>

disalah SEBUAH a::

< >(1.41)

<L L>

syarikat DISANA:

<l l> (0.93)

<l L>
A  So when I ... what's this ... attachment training in one of the firms ... a firm there.

Extract 5.16

15A  bila dah masuk dua kali salam
    dia sebutlah (0.21)
    apa tu (0.41)
    <l i> (0.51)
    <p p>
    dia punya nama: khulafak arrasyidin
    <al al> (0.29)

B  After giving two greetings, he mentioned, what's that, the name of khalifa arrasyidin.

Extract 5.17

16A  saya diberitahu:
    < > (0.95)
    bahawa apabila kita membaca
    < > (1.27)
    walaupun satu huruf (0.23) %
    > > (1.16)
kita akan dibagi: a: %

<

>(0.85)

<>

apa nama ni(0.23) %

<

>(0.3)

<L L>

<p>

sepuluh kali ganda pahalalah

<

>(0.17)

A I was told that when we read even if one letter we will be given what do you call that ten times rewards

Extract 5.18

17B jadi saya rasa (0.54)

<

>(0.94)

tujuannya ialah (0.15)

<

>(0.86)

<H H>

<l l>

supaya: (0.21)

<

>(0.59)

<

>(0.65)
mendapat (0.58)
<    >(0.49)
a:: (0.31)
<>(0.58)
-> apa yang dipanggil (0.35)
<    > (0.91)
<\L\L>
kestabilan
<    >(0.65)
< CRES >

B So I feel the purpose is to have, what is referred, to as stability.

"apa ni" (Extract 5.15), "apa tu" (Extract 5.16), "apa nama ni" (extract 5.17) and "apa yang dipanggil" (5.18) which are "wedged into" the main utterance do not only disturb the syntactic and semantic cohesion of speech but also its prosodic flow. These parenthetical segments are incomplete by themselves. In these instances, they are set off from the primary utterance by a pause and their overall low pitch.

In Extract 5.15, the speaker distinguishes the first segment "jadi masa saya" (so when I), which is syntactically and semantically incomplete, from the embedded segment "apa ni" (what's this) by a shift in speech rate and pitch. The lento tempo of the first segment gives the
impression that the speaker is facing some sort of planning difficulty. The lengthening of "ya" allows him planning time; however, the speaker still needs planning time as a result of which he produces another utterance which not only breaks the prosodic flow of the prior talk but also its syntactic and semantic flow. This inserted segment which is uttered relatively fast and begins with a step up in pitch gives him planning time as well as allows him to keep his turn. The subsequent segment i.e "attachment training" is set off from the inserted segment by a pause and a shift in pitch, i.e. with a pitch step up from low and level pitched "ni". It is also marked by a high overall pitch and forte loudness.

In Extract 5.16, "apa tu" is separated from the prior talk by a brief pause of 0.21 second and the subsequent talk by a pause of 0.42 second. There is also a step down in pitch from "lah" the last syllable of the prior talk to "a" the first syllable of the embedded segment "apa tu". The overall pitch is lower than the prior and subsequent segments.

Likewise, in 5.17 the speaker introduces a new utterance "apa nama ni" (what that is) before completing his earlier utterances and indicates a break in syntactic and semantic cohesion by demarcating the embedded segment with pauses. The overall loudness is relatively softer than the prior or subsequent talks. It is also uttered with an overall low pitch and a relatively faster tempo than the segments which bound it. The embedded
segments in Extract 5.16 "apa tu" and in Extract 5.17 "apa nama" serve the same function as "apa ni" in Extract 5.15 respectively.

Likewise in Extract 5.18, the stretch of speech is broken up by a parenthesis "apa yang dipanggil" what is referred to as whose beginning and ending are bounded by pauses. The embedded segment has a relatively slower tempo and low overall pitch. The non-low pitch at the end of the embedded segment signals incompletion.

Another type of parenthetic segment found in our data is what Nosek (1973) defines as "...a dependent satellite part of the utterance, wedged into a non-compact primary (frame) utterance from which it differs. Parenthesis ... expresses a secondary communication". Unlike the embedded segments in Extracts 5.15 through 5.18 whose purpose is to buy speaking time, the parenthetic segments in Extracts 5.19 and 5.20 are introduced for the purpose of communicating secondary information. For example, in 5.19 the speaker inserts "yalah kata orang tu" (yes like what people say), and "tak silap saya" (if I'm not mistaken) to communicate something other than what is communicated in the primary utterance. This "secondary communication", as realised in Extracts 5.19-5.20 below, is prosodically distinguished from the segmented chunks of the primary utterance:
Extract 5.19

17B yang sebenarnya saya: (1.12)

\[
< \quad > (1.46) \\
< > (0.44)
\]

\[
< H \quad H > < L \quad L > \\
< \quad \text{DEC} \quad >
\]

\[\rightarrow \text{Belayah kata orang tu::} \quad (0.74) \quad a:\]

\[
< \quad > \text{creaky}
\]

\[
< l \quad l > (1.11) < > (0.4)
\]

\[
< l \quad l >
\]

\[
< > (0.51)
\]

berminat pada dia:

\[
< l \quad l > (0.92)
\]

ataupun jatuh cinta pada dia

\[
\text{B} \quad \text{Actually I ..., yes like what people say, a: was attracted to her or fell in love with her.}
\]

Extract 5.20

18A semalam dilapurkan a: (0.43)

tak silap saya (0.39)

\[
< L \quad L >
\]

\[
< l \quad l >
\]
Timbalan Mentri Kesihatan agaknya mengumumkan akan mengadakan perang perang terhadap AIDS

A Yesterday it was reported that, if I'm not mistaken, perhaps the Deputy Minister of Health will have campaigns against AIDS.

In Extract 5.19, the segment "yang sebenarnya" (actually) is broken off from the subsequent segment "yalah kata orang tu" (yes like what people say) of the utterance by a lengthy pause of 1120 msec. The latter is often used by a speaker when he wants to qualify that whatever is said in the subsequent utterance is described by some people as such. Although "ya" the final syllable of "saya" has a level pitch which neither rises nor falls, its loudness seems to trail off until it reaches inaudibility. There is a slight pitch step-up at the beginning of "yalah kata orang tu" to indicate the beginning of a new utterance. Although the final syllable "tu" is uttered low, the pitch is level indicating that the speaker has not finished yet. Thus, although the embedded segment is separated from the subsequent talk by a pause of 740 msec., the level pitch of "tu" provides the link-up with the subsequent utterances which the prior segmented talk is part of. According to Bolinger (1986: 82) the "maintenance of the level of the final accent indicates the intention to continue".
In Extract 5.20, the parenthesis "tak silap saya" (if I'm not mistaken) which communicates how the speaker regards the primary utterance is separated from the preceding segment by 430 msec and the subsequent by a pause of 390 msec. The overall pitch is lower than the prior and subsequent segments and uttered relatively lento.

5.7 Conclusion

In the above sections, attempts have been made to identify segmented chunks of speech whose boundaries are set off by prosodic cues and the examination of such units highlights the role that they play in the ongoing discourse. Extracts 5.7-5.14 in 5.6.2., illustrate how the speaker repairs the segments which are abandoned and left incomplete by reproducing them more fluently in the second attempt. These abandoned segments lack a nucleus, yet they are well set off by prosodic cues. What is of interest here is how the prosodic cues are deployed to mark the division between performance "errors" (e.g. false starts, slips of the tongue) and self repairs, thereby providing a plausible reason for segmentation.

Likewise in section 5.6.1, an incomplete segment which is a realisation of non-fluent spontaneous speech is set off from the subsequent talk, a continuation of the former, by audible prosodic cues. The prosodic break brought about by pausing and syllable lengthening gives the speaker
planning time. Although the segments are incomplete in the sense of lacking a nucleus or being semantically and/or syntactically incomplete, the prosodic cues which demarcate their boundaries make them hearable as separate segments.

It is inevitable that in segmenting utterances into phonetically defined units, one would come across instances where boundaries of phonetic entity do not correspond with boundaries of syntactic constituent (cf. section 5.6.1 in particular). This is so because unlike the latter, phonetic entity is basically some sort of unit of performance. Pauses, for example, may break up close knit grammatical constituents or bound false starts, slips of the tongue, or incoherent segments. In so far as pauses or some other prosodic cues arise from a real need for verbal planning, their occurrence can be regarded as a direct result of the fact that speech is produced in real time and real settings. To ignore the presence of the audible cues between segments means to ignore the natural phenomena whose presence are welcome as overt indications of processing activity.