CHAPTER 4

RESULTS

4.1 TYPES OF DATA

In the study there were three types of data collected via three different methods. Firstly, the think-aloud protocols of four subjects constituted the primary data. This data which formed the basis in answering the four research questions in the study. These protocols were transcribed from the audiotaping of the subjects' articulation as they read the texts and thought aloud while completing the cloze tasks. The second type of data was obtained via stimulated recall. It comprised responses from the subjects when they were interviewed on how they arrived at their answers. The researcher noted down pertinent points as the subjects related their train of thought, particularly during their repetitive reading and pauses. The last type of data comprised relevant information from individual interview sessions with eight English language instructors. The various strategies taught by the instructors on how to solve cloze texts with options and without options were noted down during each interview session.

The procedures in processing and analysing each type of data are further explicated in the following sections. Last but not least, the results of the study are discussed in the light of answering the research questions.

4.2 PROCEDURES IN PROCESSING AND ANALYSING DATA

It was necessary to process the three types of raw data obtained before they could be analysed systematically. The following two sections give an account of how the data was processed and analysed.

4.2.1 Think-Aloud Protocol and Stimulus Recall

There was an initial attempt to analyse the think-aloud protocols as a continuous text. Since the focusing of the study is not on the overall or macro level approach to solving cloze texts (e.g. reading the entire paragraph or text before solving individual items), this analysis procedure was found to be unsystematic. This was because all the subjects retracked, that is, switched to and fro in a non-linear manner. For example, they retracked from the fifth item to any of the previous items before proceeding to the sixth item and so forth. This resulted in a rather haphazard approach taken to pinpoint the strategies employed by the subjects to solve a particular item. Hence, a more concerted alternative was implemented prior to the analysis. Firstly, all the parts specific to each item were selected from the protocol. Secondly, these parts were combined according to each item in the form of paragraphs as an indication of the chronological discontinuity. This pre-analysis procedure was administered for all the eight think-aloud transcriptions.

Next the notes from the stimulus recall were placed in a separate column and aligned with the appropriate paragraphs of the think-aloud protocols so that there would be synchronisation between the notes and the protocols. This incorporation facilitated the identification of strategies utilised by the subjects and quite often provided the necessary justification. Refer to Appendices K, K(A) and K(B), pp. 244-261. The following step was to analyse these two types of data simultaneously. Due to the subjectivity factor in analysing the data, stringent measures were taken. The data from stimulated recall were interjected at the appropriate points within each transcript. The combined data were then analysed four times for all the four subjects. This was to reduce the researcher's subjectivity when inferring the strategies employed by the subjects to solve the cloze texts. The first phase of the analysis was based on individual

protocols on cloze with options beginning with subject A and items according to numerical order from one to twenty. This was repeated with the other three subjects B, C, and D. This procedure was also conducted for the protocols on cloze without options. During this initial phase, the identification of strategies was guided by rather general classification of strategies of other researchers, such as, Tarone & Yule 1989:286, O'Malley et al. 1985a:582-584, Chamot and Kupper 1989:15&16, O'Malley and Chamot 1990:126, Black 1993:426). It also aimed to test the feasibility and suitability of these strategies for the present study. The analysis was conducted with an open mind, allowing other strategies besides those already established by other researchers. The data were examined without a predetermined list of strategies which might serve as blinkers to the researcher. The strategies were identified as and when they could be inferred from the data.

In order to eliminate subject-bias, the second phase of analysis was conducted across the four subjects, in a random order, for each item according to numerical order. During the second phase, the classification of strategies employed was found to be lacking in specificity. Consequently there were difficulties in identifying strategies that had overlapping applications. For example, "contextualisation" and *inferencing* were categorised as two discrete strategies. In order to infer, it would be necessary to examine the context. Hence, it would be more logical to label contextualisation as a superordinate for *inferencing* and other context related strategies such as, *referencing,* instead of a strategy on its own. Further refinement of the types of strategies occurred as the analysis progressed. For example, the strategy "guessing" was difficult to determine as the subjects did not state whether they were guessing the answer in any of their protocols. consequently, this strategy was replaced by *substitution* which described more accurately what the subjects were actually doing. The purpose of the

refinement exercise was aimed at delineating the various strategies so that each strategy would be distinct and could be easily distinguished from another. Strategies were identified in accordance to their definitions. Majority of these strategies were overt in nature as they described what the subjects had actually done to solve the cloze texts. The more covert strategies such as *understanding*, and *reasoning* were inferred from the context of the think-aloud protocols and stimulated recall.

Due to the fine-tuning of the types of strategies, the third phase of analysis was carried out in accordance with the changes made to the types of strategies. Similar to the first phase, the four protocols on cloze with options were analysed individually, beginning with subject A for items one to twenty. Then it was followed by the rest of the subjects. This was repeated with the four protocols on cloze without options. The last phase was conducted to maintain the accuracy of the analysis, beginning with the protocols on cloze with options followed by the protocols on cloze without options. Like the second phase, the subjects were selected in a random order. However, the items were not analysed according to numerical order but rather the items for function words were examined before those of content words. This was carried out in order to check the consistency of the analysis because the same items should warrant similar analysis regardless of the sequential order. When the analysis was finally completed after the duration of about one year, the results were tabulated. Then they were presented graphically for visual clarity (Figs. 4.1 and 4.2, pp. 267-268).

4.2.2 Interview With English Language Instructors

The strategies taught by the instructors were tabulated to facilitate the comparison between these strategies and those utilised by the subjects. The frequency of these strategies was also noted to distinguish the common strategies taught in order to verify

if these strategies were also commonly employed by the subjects (Tables 4.16-4.21, pp.278-281).

4.3 ANSWERING RESEARCH QUESTIONS

The present study seeks to provide plausible answers to four questions which examine basically the mental processes and the strategies employed by the subjects in solving rational cloze with options and without options. The basic assumption made is that every response of the subjects had resulted from their interaction with the printed text, drawing on relevant prior experience and learning as they read and completed the cloze texts. In the discussing the research questions, the name of strategies are in *italic* and 'f' symbolises 'frequency'. The words, phrases or clauses quoted from the cloze texts take the Arial font. Correct options are **bold** while distracters are ***bold***. Correct responses generated by the subjects in relation to the items are <u>**underlined**</u> while incorrect words or phrases are <u>underlined</u>.

4.3.1 Research Question 1

What are the kinds of mental processes involved in solving rational cloze with options and without options?

With reference to the auto-strategic language processing model proposed in the study, there are two possible routes which the subjects could take to select or produce a word in order to replace that which is deleted from the text. They are the automatic processing and the strategic processing of the language. There is sufficient evidence provided by the think-aloud protocols in the study to support these two kinds of mental processes.

The automatic processing is characterised by the spontaneous production of a response to fill in a blank within the text when it is read for the first time. In the think-aloud protocols, it could be recognised as stating a word to replace the deleted word without any pause during the initial reading of the text. The analysis of the think-aloud protocols in the study revealed quite a significant frequency in the occurrence of automatic language processing. (Table 4.1, p.262) illustrates the frequency of automatic language processing as demonstrated by the subjects. It can be inferred that there could be four significant factors that may induce the automatic mode of language processing. They are the types of words deleted or required to be replaced, items, cloze tasks, and subjects' preferred style of solving cloze texts. Since the study focuses on the types of the words deleted, items, and cloze tasks, the discussion that follows is limited to these three factors.

The data showed a marked difference in the frequency of the automatic mode employed between the replacement of content words and function words. This sharp contrast provided evidence to qualify the choice of word types deleted as the most dominant factor (among the three mentioned earlier) in determining the employment of the automatic mode. In addition, the data revealed that its usage was pre-eminently for producing function words rather than content words to replace the words deleted. Its fourteen occurrences in the production of function words as compared with its single occurrence in the production of content words could indicate that the production of the former triggered the automatic mode of language processing more readily than the production of content words. The function words produced by the automatic mode in the study comprised the indefinite article or determiner 'an'; auxiliary 'is'; conjunction 'and'; phrasal verbs care 'about', looking 'at'; pronouns 'it', 'what', and preposition 'in'. All these answers were correct except for the conjunction 'and'. In quantitative

terms, thirteen answers out of fifteen (that is 86.67 %) produced by the automatic mode were correct. This relatively high degree of accuracy could be interpreted as the subjects' level of comprehension of the text was precise and their English language proficiency was at the appropriate level which enabled them to produce the exact words deleted. In addition it demonstrated that when language is acquired (through exposure and regular use), language processing is via the automatic mode. The subjects appeared to be confident in the answers generated as they did not switch to any other strategic means of language processing such as *testing*, *substitution* or r*eferencing* except in two similar instances when the answers were incorrect (that is, the production of the conjunction '*and*' in the cloze without options). These errors could be a result of misrepresentation of the meaning encoded in the complex sentence. In this particular study, the employment of the automatic mode of processing did not guarantee that the correct response would be generated. This phenomenon may provide evidence to support the hypothesis that the automatic mode of language processing is absolutely a conscious procedure which draws upon cognitive processes. Hence, terms such as 'unconscious' or 'sub-conscious' used in language processing theories may be challenged and may be found inappropriate or inaccurate in depicting the processes involved.

The second factor that may exert an influence over the automatic mode of language processing could be the types of item according to their positions beginning, mid, end of the sentences and references (anaphoric, cataphoric). In the study conducted, the mid position of the deleted word in a sentence for both the cloze types was the most favourable of the three positions in setting the automatic mode into action. This was because all the answers produced via the automatic mode were words deleted in the middle and pre- or post- middle of a sentence. The non-occurrence of the automatic

mode to generate a word to fill in the gap at the beginning of a sentence could be due to the lack of context. On the other hand, the words deleted at the end of a sentence comprised content words. Consequently the task of producing a suitable word was more challenging. Hence, the strategic mode was employed instead of the automatic mode. Next, items that require anaphoric referencing seemed more conducive in mobilising the automatic mode than those requiring cataphoric referencing since the frequency was five for anaphoric referencing and three for cataphoric referencing. A logical explanation for this occurrence would be that sufficient context was provided before the blank.

The third factor which seemed to be of the least significance to induce the automatic mode might be the types of cloze (namely, cloze with options and cloze without options). This was because the data showed that the automatic mode of language processing was employed in both of the cloze types with a slight difference in its frequency which was six (that is 40%) in the cloze with options and nine (that is 60%) in the cloze without options. This difference revealed that the cloze without options lent itself to the automatic mode more readily than the cloze with options. A plausible reason for this variation in frequency could be due to a certain degree of the subjects' dependence upon the options provided which resulted in the slight reluctance to generate the answer via the automatic mode. On the other hand the existence of the automatic mode in the cloze with options revealed that self-production of responses occurred even with the providence of responses to select from. Thus it would be dubious to presuppose that in solving cloze with options, there would be no production of responses but mere selection due to the availability of the options. However this production would be contingent upon variables such as the subjects' preferred procedure or manner of solving this cloze type, their level of proficiency, the types and

location of words deleted as well as the context. The statistics drawn from the data indicated that the automatic mode comprised only 9.38% (that is 15 responses out of 160 responses) whereas the remaining responses were solved using the strategic mode. This vast majority for the strategic mode could be a consequence of the subjects' cautious style of solving the cloze exercise. In other words, they would prefer to read the sentences or related sentences before producing or selecting a response. In doing so, they might be suppressing or unwilling to activate the automatic mode.

The data obtained in the study revealed that the automatic mode of language processing is a dynamic and conscious process. It is more frequently operative in the cloze without options than the cloze with options. In addition the automatic mode is utilised to produce function words rather than content words which are preferably deleted in mid-sentence and require anaphoric referencing. The specific strategies which are employed for the strategic mode will be discussed in further details in the findings related to the second research question.

4.3.2 Research Question 2

What strategies did the subjects use to solve rational cloze

(a) with options; and

(b) without options?

The analysis of the think-aloud protocols revealed that all the four subjects had utilised a wide range of strategies in their attempt to solve the two types of cloze texts. A total of thirty-four strategies had been identified. They had employed "clusters" (6 to 12 types) of strategies in solving each item. The description of strategies utilised by the subjects would be based upon how frequent these strategies occurred. The frequency for each strategy was the total number of times it was employed by all the four subjects

when solving the ten content or function words in each type of cloze text. Table 4.2, p.262 gives the range of frequencies referred to in this discussion. The average frequency scale was based on the average frequency of the four types of categories involved (namely, (i) content words with options, (ii) function words with options, (iii) content words without options, and (iv) function words without options) with a range of 16 to 20.

The response to this inquiry would be viewed from two perspectives, that is, the strategies employed to solve rational cloze with options and rational cloze without options. For each type of cloze, there would be a comparison made between the strategies utilised to solve content words and function words.

4.3.2.1 Strategies employed in solving cloze with options

The strategies utilised for solving the content words will be elucidated followed by those for solving the function words. Then a comparison will be made between the strategies utilised for the content words and the function words.

4.3.2.1 (a) <u>Content words with options</u>

With reference to the data in Table 4.3, p.263, the most frequent strategy employed by the subjects to solve the content words was *pre-deletion and post-deletion* (f = 177). The subjects had opted to reread word, phrases, clauses or sentences before and after the blank as a strategic move to increase their level of comprehension of the particular context circumscribing the blank concerned. This strategy was by far the most commonly employed by all the four subjects to solve each of the content word item with an average of 17 or 18 occurrences per content word. The next two strategies within this **ultra high** frequency range but with approximately half the frequency were *testing* and *substitution* (f = 90 each) followed by *elimination* (f = 47). These strategies

136

were observed to be used in clusters, that is, as a group. To a considerable extent, they occurred in a regular chronological order, beginning with *substitution*, followed by *testing*, and lastly, *elimination*. The probable reason for the frequency of *elimination* to be approximately half of *substitution* and *testing* could be that this particular strategy was not verbalised as often or inferred as vividly as *substitution* and *testing*. The availability of four options for each item could be the major contributing factor for the frequent employment of these strategies by the subjects. The data also revealed that these three strategies were usually intervened by *pre-deletion and post-deletion*. There was only one strategy that was observed within the **very high** frequency range. Subjects placed emphasis on the meaning of words presented as options. *Semantics* (f = 32) was aptly utilised by subjects to select the appropriate content words based on their meaning. This particular strategy was restricted to only six items; each with options comprising four different words. Subjects did not utilise this strategy to gauge the answers for the remaining four items because the options were derivatives of one particular word. For example, options for item 3 were communicate, communication, communicative, and communicatively. For these items with options consisting of derivatives of a particular word, subjects had utilised a more appropriate strategy, that is, *morphology*. Next in the **high** frequency range, the strategies employed were *equivalent* (f = 29), *reasoning* (f = 29), and *response monitoring* (f = 26). The subjects exercised *equivalent* to enhance their comprehension of the text as well as the meaning of some options. It was inferred that the subjects' route taken to arrive at their answers was based upon *reasoning* with the exception of items 3 and 10. For these two items, subjects exploited the text by referring to the words stated in the text. Hence, they utilised *stated* more prominently than *reasoning* where these two items were concerned. The subjects also monitored their responses to ensure the certainty of their choice of words in completing a particular portion of the

137

text. There were two items, 4 and 5, which all the subjects were perceived to be uncertain of their response. They were noted to be substituting and testing the options even during their final attempt to arrive at what they would consider a suitable response. Consequently, they were inferred as not being able to monitor their responses for these two items. On the other hand a plausible explanation would be that the monitoring of responses could have occurred simultaneously with the final *testing*.

The strategies within the **above average** frequency range were *referencing* (f = 24) and *expression* (f = 21). The subjects had applied *referencing* in their response to items 3 and 8 as both the responses were stated within the text. The subjects also exploited this strategy by referring to the words or phrases before or after a particular item to establish the context prior to selecting their response. *Expression* was perceived to be utilised in solving only four items, namely, 3, 6, 16, and 19. These items rendered themselves to this strategy due to common occurrences of expressions, such as, "**communication process**", and "**prepare a proposal**". Three strategies were noted to exist within the **average** frequency range, namely, *inferencing* (f = 18), *deduction* (f = 18), and *morphology* (f = 17). After having established the context, subjects made their own inferences before choosing their responses. For example, from the textual evidence that there were two persons talking, subjects B and D inferred that a conversation was taking place. Hence, they selected the option "**conversation**" for item 1 correctly. The subjects could have employed the same strategy in solving items 4 and 17. Both these items required them to infer the emotion of the person concerned. Instead they paid more attention to choosing which option could fit the group of "**misunderstood, confused**". Next the subjects were inferred to have utilised *deduction* and *morphology* to solve the four items with derivatives as options. They had used their grammatical knowledge in their attempt to select the most suitable response for each item. Besides

these items, subjects A, B, and C also noticed inflexions such as 'ing', 'ed' and 's' of certain words in the text and the options. A probable reason for the frequency of these strategies to be noted within the average range could be the limited opportunity made available in this cloze exercise. The only strategy within the **below average** frequency range was perceived to be *delayed response* (f = 14). This strategy was observed to be utilised by all the subjects for different items except for items 2 and 4. As such, this strategy could be considered as a cautious move practised by the subjects when they were uncertain of their responses.

Within the **low** frequency range, subjects were observed to exercise *association* (f = 9), *grouping* (f = 7), and *analysis* (f = 6). All the subjects employed *association* in solving item 16 whereby they utilised their schema of movies or world knowledge to gauge the most appropriate response. Only subject B was inferred as employing this strategy for items 3 and 8 to associate verb with process, and reaction with action respectively based on schematic knowledge. *Grouping* was utilised by all the subjects for item 4. Subjects B and D were noted to have employ *analysis* for item 4 whereas subject B was observed to have exercised this strategy for item 17. These subjects had raised questions concerning the meaning of certain options, namely, "**frustrated**", and "***pitiful***". Subject C remarked that the particular sentence with item 4 was very long while subject D wondered aloud which option the answer should be. This strategy was recorded as possessing a low frequency most probably because the subjects had not verbalised their doubts and comments about the text or the task. Another plausible explanation would be that the subjects were familiar with the task and they were concentrating solely on solving the items in the cloze text.

There were eight strategies observed within the **very low** frequency range. The four more frequent strategies included *stated* (f = 5), *induction* (f = 4), *word retrieval* (f = 3), and *auditory monitoring* (f = 3). The other four strategies with the frequency of one were *linking answers, phonology, language transfer,* and *alternative response*. The subjects employed *stated* to solve items 3, and 8. Subjects A and C selected "**communication**" as the answer for item 3 as this word was mentioned in the text quite frequently. Likewise the word "reaction" which was stated in the sentence after item 2 had led subjects A, B and C to choose the option "**reaction**" for item 8. *Induction* was utilised by three subjects for four different items comprising two items with vocabulary as options and two items with derivatives as options. Subject A employed this strategy twice. Firstly, she noted that two words, "specialising" and "facilitating" located before item 5 bore the common inflexion, that is, "ing". She deduced that the answer would also end in "ing". The same subject also identified the use of the present simple tense from the word "suggest" before item 19. Subject B noticed the use of "ed" in the word "confused" before item 4 and in two options, "*stressed*" and "**frustrated**". She narrowed down the answers to these two options. Subject C observed the use of the present simple tense in the paragraph encompassing item 6 and deduced that the answer would also be in the same tense. Next *word retrieval* was employed by only two subjects, A and D in solving item 19. Subject A suggested two words, "<u>one</u>" and "<u>other</u>" as alternative answers while subject D suggested only one word, that is "<u>the</u>". The occurrence of this strategy was rather unexpected as there was no need for the subjects to generate their own words. In fact it would be against the nature of this task and subjects would be penalised for doing so. Apparently the subjects were uncertain of the suitability of the four options given and they were actually attempting to suggest their own answer and then select the option which would be synonymous with their answer. This strategy could have been utilised to resolve their uncertainties about the

140

meaning and usage of the four options. The other strategy with the same frequency of three was *auditory monitoring*. It was utilised by subject A for two items, 3 and 19, and subject D for item 5. The subjects evaluated the suitability of the options in completing each of these items. In all three instances, the result of using this strategy had led to elimination of the options concerned on the basis that they did not sound right. The subjects seemed to be dubious about which option should be the most suitable response. The use of this strategy appeared to have assisted the subjects in deciding which options could be considered as inappropriate.

In addition there were four strategies that were inferred to be employed only once by three subjects to help solve three different items. Firstly, subject A utilised *linking answers* for item 3 when the response for item 2, that is, "**conversation**" was noted to possess the suffix "tion". The use of this suffix guided subject A to select the option "**communication**" as the response for item 3. Secondly, subject A also employed *alternative response* for item 19 when she generated two other words of her own, "one" and "other", and tested them one after the other. Thirdly, *phonology* was utilised by subject B to gauge the meaning of the word "*pitiful*" as the sound of "pity" reflected a "sad" emotion. Fourthly, subject C applied *language transfer* in choosing the option "*future*" for item 19. The subject's knowledge of the Mandarin expression "*future* days" had influenced the selection o the option "*future*" as the response. All these strategies except for *alternative response* were drawn from the stimulated recall data. This could point to a possible factor that contributed to the very low frequency of these three strategies, that is, they were not verbalised as often as they were actually utilised.

The last category consisting of nine strategies to be examined would be those that were considered to be with zero frequency due to being **"absent", "non-verbalised" and/or**

"uninferred". There were five strategies listed in this category which could be regarded as "absent". These included *focusing, group retrieval, production modulating, collocation,* and *syntax*. The first strategy was most probably undetected due to the subjects' familiarity with the nature of the task. Hence, they knew exactly what they were required to do and they did not need to ask any questions concerning the task or make any comments about the task. The next two strategies, that is, *group retrieval,* and *production modulating* might not be utilised because of the type of cloze. Since the options were given for each item, there was no necessity to retrieve group of related words, producing own words, and modulating them. Furthermore, there were no observable attempts made to replace the deleted words instantaneously as the text was read due to the subjects' reliance upon the options provided. Next, *collocation* could not be observed as a result of the nature of the words deleted. None of these words were collocates of the word before or after the blank. The subjects' attention was noted to be so focused upon the options that they did not seem to remember that they could have also examined the syntactic nature of the sentences in the text. Although the last four strategies, that is, *interpreting, synonym, understanding,* and *comprehension monitoring* could not be inferred directly due to non-verbalisation, they could have been utilised while the subjects were reading the text repeatedly. These strategies could have assisted the processing of the information or message of the text. As they might have occurred when the text was read aloud, the subjects did not verbalise how these strategies were utilised, that is, what went on in their mind while reading aloud or during rare moments of short pauses. Hence, one of the drawbacks of think-aloud protocol could be evident at this point of data analysis. On the other hand, data from the stimulated recall recorded a general statement from the subjects that they were actually actively employing these four strategies when they were reading aloud or during their pauses. Therefore the "zero" frequency assigned to these strategies could have been the

consequence of non-verbalisation or failure of direct inference based on the think-aloud protocols and not their absence per se.

4.3.2.1 (b) Function words with options

With reference to Table 4.4, p.264, there were five strategies observed within the **ultra high** frequency range. The most frequently utilised strategy was *pre-deletion and post-deletion* (f = 171) which was utilised by the four subjects in solving all ten function words deleted. This was followed by *testing* (f = 90) and *substitution* (f = 67). All the subjects were observed to employ *testing* to a great extent in order to solve the majority of the ten items. There were instances when subjects did not seem to utilise *testing*. For example, subject A did not test the answer selected for items 1, 9 and 12. Likewise subject C was not observed to have employed *testing* for items 1 and 20. These subjects exhibited confidence in their choice of answers and it was unnecessary for them to test them. *Substitution* was noted to be utilised by the subjects for all items except item 1. They seemed satisfied with their response for this particular item. The subjects would normally substitute each of the options in turn and test them for their suitability. From the frequency obtained, it was evident that the subjects tested each option more often than substituted the options. In other words, each *substitution* could be followed by more than one *testing*. The remaining two strategies within the ultra high frequency range were *expression* (f = 42) and *elimination* (f = 37). The subjects employed *expression* for five items. It was utilised more frequently for items 7, 11, and 14 than for items 9 and 18. The former set of items involved the use of preposition, "available in" and phrasal verbs, "look for" and "care about". The subjects utilised their knowledge of common expressions, that is, the subtle differences in meaning and usage in order to select the best option. The latter set of items was concerned with semantic markers which constituted the options. *Expression* was utilised in two out of three items

with semantic markers deleted. Firstly, it was used only once by subjects A and D for item 9 with the expression, to "**see if**". Secondly, subjects A, B, and C employed it once for item 18 with the expression "*in contrast*". However, this was a wrong expression. Since these semantic markers exist as expressions the subjects had aptly utilised this strategy. The subjects were not able to do likewise for the item 13 although it also involved semantic markers. This was because the deleted word "**But**" does not exist as an expression. The next strategy, *elimination*, was manifested in solving eight out of ten items. For each of these items, there were only two or three subjects who were observed to have employed this strategy. This strategy was noted to be utilised after *testing*, particularly when the subjects seemed to be uncertain as to which option to choose. They eliminated the unlikely options in order to identify the most suitable response. They were quite certain of their response for item 1 on pronouns and item 15 on auxiliaries. Its low ranking within the ultra high frequency range could be due to non-verbalisation and/or the subjects would not eliminate an option until much *testing* was deployed to determine the suitability of the options.

There was only one strategy inferred within the **very high** frequency range, that is, *deduction*. The subjects applied their knowledge of grammatical rules to assist them in selecting their response. For example, with options comprising semantic markers, such as, "*And,* **But,** *Thus,* **In addition**" in item 13, the subjects were able to state their function. In another instance, they chose the indefinite article or determiner "**a**" for item 20 because the positive result stated was not definite and the immediate word after the deletion, "positive" began with a consonant.

There were two strategies inferred within the **high** frequency range. These were *reasoning* (f = 30) and *response monitoring* (f = 27). All the subjects were noted to

utilise *reasoning* in assisting them to solve items related to pronouns, (item 1), semantic markers (item 13), and auxiliaries (item 15). *Reasoning* was also employed by two or three subjects to select the responses for the remaining seven items. Subjects monitored their responses for all the items except item 20. This could be due to their certainty of their choice of answers.

In the **above average** frequency range, *equivalent* (f = 22) and *referencing* (f = 21) were observed to be applied by subjects in solving certain items. With reference to the former strategy, there were two languages utilised, that is, the Malay language and Mandarin at the word, phrase and clause levels. All the subjects employed this strategy in their attempt to solve item 7 which comprised prepositions as options. They appeared to be quite uncertain of the suitability of the options as each option seemed to be meaningful within the context of the sentence. They translated the options into the Malay language in order to test or confirm the suitability of each of the options. This was also manifested in all items involving semantic markers as options (that is, items 9, 13 and 18). Only subject A was observed to have employed Mandarin for items 7 and 15. It was used to translate **"produce results"** and **"available"** in item 7 as well as **"you don't know"** in item 15 in order to enhance comprehension of the text. The latter strategy *referencing*, was deployed by all four subjects in solving item 12 with relative pronouns as options, item 15 with auxiliaries as options, and item 20 with determiners as options. They referred to the text for clues in order to select the most appropriate answer. For example, they referred to the phrase **"positive result"** to determine that the answer should be **"a"** for item 20.

There were no strategies observed or inferred within the **average** frequency range. This phenomenon implied that the frequency of the strategies utilised was distributed above

and below the average frequency range. The strategies within the **below average** frequency range were *inferencing* (f = 14) and *delayed response* (f = 13). It was noted that *inferencing* was employed for five items; twice for options that were semantic markers (that is, items 13 and 18) and once for options that were pronouns, relative pronouns, and auxiliaries (that is, items 1, 12 and 15). All four subjects inferred that a person should say "something" and the relative pronoun for "something" would be "**what**" for item 12. Another example, subject A was able to infer that "What you don't know" is a singular matter. *Delayed response* was utilised by all the subjects except subject B for five items (that is, items 7, 9, 15, 18 and 20). This strategy was manifested when subjects retracked in their reading without any responses. It was observed to occur in the middle of their attempt when they were uncertain as to which response was the most suitable. This strategy allowed them to enhance their comprehension of the context surrounding the deleted word before they could select the response.

In the **low** frequency range, there was only one strategy recorded, that is, *semantics* (f = 8). Three subjects, A, B and C, employed this strategy in solving four items. These included all the three items with prepositions as options (that is, items 7, 11, and 14) as well as one item with semantic markers as options (that is, item 13). All the three subjects were attempting to differentiate the meaning of the phrasal verbs e.g. between "care *of*" and "care **about**", "look **for**" and "look *at*". Subject A tried to gauge the meaning of "eager" and "available" from the context but was incorrect. The employment of this strategy could have occurred for all other items but was not verbalised. All the subjects could have exercised this strategy inaudibly while reading the text and the options aloud. On the other hand, the low frequency for this strategy could provide evidence that comprehension of text would more likely to occur in

chunks than at word level. *Semantics* was only employed when the need arose as revealed by the data.

Within the **very low** frequency range, there were twelve strategies recorded. To begin with, *induction* (f = 4), was utilised for only item 15 with (auxiliaries as options) by three subjects. They were able to select words from the sentence with the item that would assist them in identifying the tense. As an illustration, subjects B and D, recognised the present simple tense from the words "don't know" and concluded that the suitable auxiliary should also be in the present simple tense. The next three strategies, namely, *interpreting*, *analysis*, and *language transfer* were observed to have the same frequency of three each. Firstly, *interpreting* was observed for two items (that is, items 9 and 13). Both these items were concerned with semantic markers as options. Subject B was able to interpret from the text that there was neither result nor cause stated before or after the deleted word for item 9. Subject C could interpret that the sentence before item 13 "sorry, I don't have time to see you now" as a negative response to a request made earlier. Secondly, *analysis* was observed to be subject specific as it was utilised by only subject A in enhancing comprehension of the context for items 14 and 20. The analysis was realised through questions asked by the subject pertaining to context and word meaning. The respective questions were "What does he say?" and "What is the meaning of accomplish?" Thirdly, *language transfer* was observed to be utilised by two subjects, C and D, for only one item, that is, item 14. Both subjects had selected the option "*of*" as the response based on their prior knowledge of the expression "take care of". This could be considered as negative language transfer due to overgeneralization of the use of this expression in the second language. Strategies with a frequency of two included *collocation*, *syntax*, *word retrieval* and *auditory monitoring*. Subjects A and C were observed to utilise

147

collocation, that is, "**instead of**" for item 18. Subjects B and C were inferred to employ *syntax* for item 1 when they noted the use of "**we**" in the beginning of a sentence and "**us**" in the middle of a sentence. These subjects had formed their own guideline on the use of "*we*" and "**us**" based on syntax and not the normal use of "we" as "subject" and "us" as "object" rule. Subject A utilised *word retrieval* when the word "<u>on</u>" was produced as an additional option for item 7 while subject B produced the word "<u>with</u>" for item 11. *Auditory monitoring* was utilised by subject A for item 14 and subject B for item 18. They gauged their choice of response by listening to the flow of the response and how frequently they had heard it being mentioned. There were three strategies that were perceived to be utilised only once. They comprised *morphology*, *comprehension monitoring*, and *focusing*. Subject A was inferred to have utilised *morphology* by changing the form of the pronoun "we" to "us". Subject A was perceived to monitor her comprehension for part of a sentence when she commented she understood that particular portion of the sentence related to item 7. She was also noted to focusing on the task requirement when she asked herself why "**is**" was the response for item 15.

There were altogether nine strategies that were absent, namely, *linking answers*, *phonology*, *grouping*, *group retrieval*, *synonym*, *alternative response*, *stated*, *production modulating*, and *understanding*. Their absence could be due to several plausible factors. Firstly, the choice of words deleted and the options were not conducive for the first six strategies mentioned. Thus the need to employ them did not arise. Secondly, the task did not require *alternative response* and *production modulating* as the options were given. The third plausible explanation could be that there was no verbalisation of these strategies, especially *understanding*. As a result these strategies were not being inferred.

4.3.2.2 Strategies employed in solving cloze without options

Similar to 4.2.2(a), the strategies utilised for solving content words would be discussed followed by function words. Then a comparison would be made between the strategies employed in solving these two types of word deletions.

4.3.2.2 (a) Content words without options

With reference to Table 4.5, p.265, there were eight strategies recorded within the **ultra high** frequency range. The most frequent strategies was *pre-deletion and post-deletion* (f = 179). It was utilised by all the subjects in helping them to understand the context for the majority of the ten items. There were certain items where subjects A (items 13 and 16), B (items 1, 6, and 16) and D (item 16) did not employ this strategy. In these instances, they were actively referring to specific words in the text that were suitable responses. As such there was no retracking of phrases or sentences in their reading. The next strategy, *testing* (f = 103), was inferred to be exercised by all four subjects for six items. Subject B did not employ this strategy for items 6 and 14 while subject C did not utilise it for item 13. Both these subjects were certain of their initial responses and did not generate any other options besides those. This strategy was utilised once by subject D for item 16. She was a little hesitant over her response and was inferred to be testing it when she repeated her response. Thirdly, *production modulating* (f = 61) was perceived to be utilised for all items by the subjects. However it was observed that only subject B did not employ this strategy for three items, namely, items 6, 9, and 16. On these occasions, subject B was quite certain of the response produced, especially for items 6 and 16. Hence there was no need to modulate the response generated. Fourthly, *referencing* (f = 58) was noted to be employed by all the subjects in assisting them to arrive at their responses for five items (that is, items 1, 6, 14, 16, and 17). The responses for these items were content words that appeared in the text. Thus subjects

149

could deduce the suitable response if they had comprehended the message of the text they were reading and made the correct references. For the remaining five items, *referencing* was employed but not by all the subjects for any one item simultaneously. Out of these five items, four responses were content words that did not appear in the text. Subject B was most consistent in utilizing *referencing* for all the five items except item 11. The subjects referred to particular words in the text and attempted to fill in the blanks with them. They also employed *alternative response* (f = 57) and created their own options. They generated more than one response as alternative responses for all items except item 16. This was because the response for this item, "listening", was very obvious and all four subjects did not doubt their response. Besides *referencing*, subjects exercised another means of generating responses, that is, *word retrieval* (f = 56). They retrieved words which were not found in the text as their responses for all ten items except item 16. It appeared that they had retrieved these words from their "mental storage" of the second language. However there were a few instances when the words were retrieved in the subjects' first language (Mandarin : Mandarin) and also in the Malay language (Bahasa Malaysia). For example, subject C retrieved the word "conversation" in Mandarin for item 1 before translating the word into the English language. Subject A retrieved the Malay word "tumpuan" which was then translated as "concentration". Furthermore the text and the context surrounding the item might have triggered this retrieval. This could have also resulted from *inferencing* (f = 41) whereby the subjects attempted to read between the lines. With reference to item 20, subject D drew a conclusion that one needs to "work hard" in order to make one's communication more precise. This strategy was inferred to be utilised by nearly all the four subjects to solve all items except for item 16 which was an item that was purely based on *referencing*. This was also observed for *equivalent* (f = 36), that is, the last strategy listed within the ultra high frequency range. Subjects A and C had utilised this strategy

for seven items; subject B five items and subject D three items. Subject B used only the Malay language while subject C used only Mandarin. The other two subjects, A and D, utilised both these languages with Mandarin as the more dominant one. There were instances when the subjects retrieved a word as a response to an item in one of the non-second languages and then translated it into the English language. For example, subject A retrieved the word "how" in Mandarin and translated it into the second language as the response for item 9. There were also instances where the reverse occurred. The subjects retrieved the response in the English language and then translated it into one of the non-second languages. For example, subject B retrieved words like "<u>action</u>" and "<u>respond</u>" for item 2 and then translated these words into the Malay language, that is, "perbuatan" and "tindak balas" respectively. Subject B explained that the translation was necessary to countercheck the suitability of the response. The third type of occurrence whereby *equivalent* was utilised by the four subjects was the literal translation of the words, expressions, phrases or clauses (parts of sentences) into the Malay language or Mandarin. The subjects commented that this was done to enhance their understanding of that particular portion of the text.

There was only one strategy observed within the **very high** frequency range, namely, *delayed response* (f = 33). Subjects A and D employed this strategy more frequently than subjects B and C. There could be a possibility that this strategy seemed to be subject specific. Subject D utilised this strategy for seven out of ten items whereas subject B utilised it only once. Subject D could be more cautious while subject B could be more confident with her responses. Another factor that influenced the employment of *delayed response* could be the level of difficulty of the item. All the subjects did not respond to item 11 immediately as the context did not provide clear indication of what the response could be. However they responded to items 14 and 16 without any

hesitation as the response to each of these items could be referred to in the text. There were no strategies observed or inferred within the **high** frequency range. This indicated that the distribution of the strategies employed was not uniform within the high, very high and ultra high frequency ranges.

Within the **above average** frequency range, the subjects were inferred to have employed *reasoning* (f = 24), *interpreting* (f = 21) and *eliminating* (f = 21). *Reasoning* was inferred to be utilised for nine out of ten items by at least one subject. The subjects decided on their responses based on logical thinking which was to a large extent influenced by their comprehension and interpretation of the context related to each item. Subject A seemed to capitalize on *interpreting* as this strategy was observed to be employed for seven out of ten items. The other three subjects were noted to employ this strategy rather sparingly, that is, for only one or two items. The subjects interpreted the meaning beyond the lines using the second language and Mandarin. They were noted to be reading certain portions of the text, particularly of unfamiliar phrases or long sentences, for their meaning. For example, subject A interpreted "communication as a 50-50 process" by equating it as "the speaker listens 50% and speaks 50% during a communication process". This strategy appeared to have the effect of assisting the subjects to comprehend new concepts, ideas or messages encoded in unfamiliar terms or complicated sentences. It functioned to break down the barrier to comprehension and make the text "digestible" or easier to be understood. As for *eliminating*, all the subjects were observed to eliminate their own responses generated for all items except items 16 and 17. Upon testing the responses generated, the subjects displayed the ability to gauge the suitability of these responses and eliminated the improbable ones. When compared with the frequency of 57 for *alternative response*,

the frequency of 21 for *eliminating* appeared to be much lower than expected. This could be due to non-verbalisation of mental elimination of the responses generated.

Next *expression* (f = 18) was one of the two strategies noted within the **average** frequency range. It was utilised for 50 percent of the items. Subjects were observed to utilise familiar expressions to select the most appropriate response and eliminate the unlikely ones. These include "<u>practise</u> your", "**think** about", "<u>learn</u> about", "passive <u>way</u>", "another <u>way</u>", and others. The subjects chose their responses based on what they perceived to be the most common combination for these expressions should be. The other strategy observed within the average frequency range was *stated* (f = 16). It was utilised for seven out of ten items by all the subjects with subject A employing it for five items; subjects C and D four items; and subject B three items. There was only one item (16), whereby all the four subjects referred to the phrase "background listening" which was stated in the text to arrive at the response "**listening**". In another instance, subject C correctly referred to the word "**exercise**" stated in the text for item 17.

There were two strategies observed within the **below average** frequency range. These were *substitution* (f = 13), and *association* (f = 11). Although the frequency for *testing* was ultra high, the frequency for *substitution* was below average. This rather unexpected observation could be due to the limited number of responses generated and the subjects were testing the same responses more frequently than substituting them with different ones. Furthermore, some responses were tested and eliminated without further substitution. The second strategy, *association*, was employed for item 6 by all subjects as well as items 14 and 17 by Subjects A and C respectively. As an example, "talking to" was associated with "a **person**" whereas "talking about" with "a topic".

Within the **low** frequency range, there were two strategies observed, namely, *synonym* (f = 8), and *auditory monitoring* (f = 7). It was inferred that only three subjects, namely, subjects A, C, and D, employed *synonym* for 50 percent of the items. Subject D utilised this strategy the most frequently, that is, for three items (1, 6, and 11). For example, "talk" was similar to "conversation" and "subject" was similar to "topic" in meaning. The subjects generated words which were similar in meaning for an item and then they would select the more appropriate word as the response by referring to the context. It was inferred that the three subjects (A, C, and D) employed *auditory monitoring* for five items (9, 11, 13, 14, and 20). Again it was subject D who utilised this strategy most frequently for four items (9, 11, 13, and 20) while subjects A and C utilised it only once for items 13 and 14 respectively. These subjects decided the suitability of their responses based on how "correctly" they sounded when their responses were read aloud within the context of the clause or sentence as a whole. For example, subject A selected "quietly" for item 13 because it "sounded more correct" than "quiet". Similarly, subject B chose "practise" instead of "practical" for item 14 because it sounded "correct".

Within the **very low** frequency range, there were seven strategies recorded. The first three with the same frequency of 5 were *deduction, semantics,* and *group retrieval.* There were only three subjects, namely, A, B, and C who exercised *deduction* for three different items. For instance, subject C was inferred to select the singular form of "word" for item 11 instead of the plural since "the last" referred to only one word. *Semantics* was observed to be utilised by two subjects, A and B, for items 9 and 2 respectively. As an illustration, subject A differentiated the meaning between "how we communicate" and "how to communicate" by equating the former with "others are talking to us" while the latter as "we are talking to others". All the subjects except subject B were observed to have employed *group retrieval* for 50% of the items. These

154

subjects generated responses that were belong to the same word class. For example, subject C produced a series of wh-word, namely, "<u>how</u>", "<u>who</u>", and "<u>what</u>" continuously for item 9. The next two strategies inferred with a frequency of 3 each were *analysis* and *morphology*. There were only two subjects, A and D, who utilised *analysis* for three items (11, 14 and 20). They raised questions such as what word would be suitable as the response, how to spell the response proposed and the meaning of certain individual words in the text. In doing so, they broke down the task into manageable parts. The low frequency of this strategy could be due to non-verbalisation. The subjects could be actually employing this strategy while they were reading aloud. Next there were only two subjects, A and C who utilised *morphology* for two items (11 and 14). For example, subject C was able to distinguish the usage of the morpheme 's' to form the plural of "word" for item 11 and the morpheme "al" to form the adjective of "practise" for item 14. Subject C considered both these forms to be inappropriate. Firstly, the context for item 11 required the singular form of "word" as in the expression "the last <u>word</u>". Secondly, the adjective "<u>practical</u>" did not sound correct in the expression "much <u>practical</u> is meaningless" to subject C. The next two strategies with a frequency of two were *focusing* and *language transfer*. Subjects A and D attempted to solve the items 6 and 1 respectively by focusing on the criterion for the response, that is, it should be limited to only one word. *Language transfer* was utilised by only subject C for items 1 and 9. The transfer of lexical meaning was from the second language to the native language and back to the second language as the response. As an illustration, subject C translated the meaning of the word "communication" into Mandarin as "conversation" which was translated into the second language as "conversation". During the *language transfer*, the translated word in the native language was again converted into the second language by using another word synonymous with the original word.

There were nine strategies which were **absent, non-verbalised and/or uninferred**. The subjects were observed not to have employed *linking answers, induction, collocation, phonology, grouping* and *auditory retrieval*. *Understanding* and *comprehension monitoring* could have been utilised while the subjects were reading the text aloud. consequently, they could not be verbalised simultaneously although they should occur before responses could be generated. Lastly, there was no *response monitoring* because there were no responses available for the subjects to select from. They were checking their self-generated responses via *production modulating* instead.

4.3.2.2 (b) <u>Function words without options</u>

With reference to Table 4.6, p.266, there were 27 strategies out of 34 that were utilised by the subjects to solve the ten function word items for cloze without options. There were six strategies inferred within the **ultra high** frequency range. These included *pre-deletion and post-deletion* (f = 141), *testing* (f = 58), *production modulating* (f = 48), *word retrieval* (f = 41), *equivalent* (f = 39), and *reasoning* (f = 38). Pre-*deletion and post-deletion*, being the most frequently utilised strategy, was essential in aiding the subjects to establish the context surrounding the item and enhance their comprehension of that particular part of the text. However there were seven occasions where this strategy was not utilized. For example, subject C was focusing solely on wanting to know the meaning of "analogy" in item 3. Hence, there was no reading of the text before or after the item. In another instance, both subjects, B and C, were very certain of their response for item 7. consequently, there was no necessity for them to reread the text adjacent to the item. About 70% of the responses produced were tested by the subjects in gauging the suitability of these responses. Subject B appeared to have utilised *testing* the least as she had employed it for four out of ten items. This could most probably be equated to the least number of responses generated, that is, only one

response each for six items. In other words, *testing* was considered to be necessary when the responses generated were to be evaluated if they were satisfactory. On the other hand, subject A utilised *testing* for nine out of ten items. The same response generated was repeatedly read together with the related portion of the text to evaluate how appropriate it was within that particular context. Another strategy that closely resembles *testing* would be *production modulating*. Along similar line of reasoning, subjects A and D were observed to have employed this strategy for nine out of ten items. This phenomenon could be attributed to their cautious manner of selecting the most appropriate answer. Checking their preferred response might also help them ascertain what the deleted word was. Subjects B and C were more were more satisfied with the form of their responses and did not need to change them often. Therefore they exercised *production modulation* less frequently. The next strategy, *word retrieval*, was utilised by the subjects for all the items. The subjects were able to generate words that were not in the text. Examples of these words were "<u>act</u>", "<u>actually</u>", "<u>learn</u>", "<u>try</u>", "<u>dialogue</u>", "<u>practise</u>", "<u>way</u>", "<u>title</u>" and "<u>inside</u>". They appeared to be generated from the subjects' lexical storage. This strategy could be subjected to the degree of comprehension of the context surrounding each item. The fifth strategy within this category was *equivalent*. It was utilized by most subjects for all ten items. They were noted to translate the response generated in the second language into the Malay language or Mandarin as a means to confirm its suitability. Subject A, in particular, was observed to translate phrases from the text into the Malay language to enhance comprehension of the text. Lastly, *reasoning* was inferred to be employed by the subjects in order to arrive at their responses. There was only one item (18), whereby the subjects, A and B, did not utilize this strategy. Instead they had referred to the words in the immediate sentences before and after the item. There were no strategies observed for the following two frequency ranges. They were **very high** and **high** frequency

ranges. In other words, there were no strategies observed or inferred between the frequency range of 26 and 35. This disproportion of the distribution in the frequency of the strategies employed revealed that only six strategies within the ultra high frequency range were repeatedly exploited by the subjects, while the remaining twenty-one strategies were utilized to a lesser degree (that is, those in the average and low frequency ranges).

The subjects were observed to exercise five strategies within the **above average** frequency range. These were *alternative response* (f = 25), *deduction* (f = 24), *inferencing* (f = 23), *delayed response* (f = 22), and *referencing* (f = 21). When the subjects doubted the suitability of their initial response, they generated other responses. There were two items (7 and 10) which none of the subjects had produced any alternative responses. They were certain of the suitability of the response they generated. Secondly, *deduction* was utilized for all the items, except items 10 and 12. The subjects applied certain grammatical rules to pinpoint the most suitable response. For example, subjects B and C selected the article or determiner "**an**" for item 3 because "**analogy**" begins with a vowel sound and it is singular. However, there were instances when inaccurate *deduction* was exercised. An example was subject C's misperception that "if" was used to link opposite ideas. There were two items, 10 and 12, which no *deduction* was utilised by any of the subjects. The responses for these two items were prepositions, 'in' and 'at', because the focusing was on meaning and not grammatical rules. Thirdly, *inferencing* was inferred to be employed by the subjects when they attempted to read in between the lines for the presupposed meaning that would assist them in identifying the appropriate response. For instance, subject B inferred that words could be generated from an idea for item 18. However, the response "idea" was inappropriate within the context of the sentence concerned. In another

instance, subject B was able to eliminate "<u>then</u>" as the response for item 19 since she inferred the action of "using a different tone" was not sequential to the prior action, that is, "suggesting something quite different". Fourthly, *delayed response* was exhibited by the subjects when they failed to generate any responses after reading the sentence with the deleted word. All the subjects exercised this strategy for item 15 while an average of two subjects utilised this strategy for five other items, that is, 3, 5, 8, 10 and 18. There were four items which the subjects responded to without delay, namely, items 4, 7, 12 and 19. The responses for these items could be considered as rather direct. In addition the sentences with these items were relatively shorter. On the other hand, the subjects required a longer time to process sentences that were more complex in nature. This could lead to a delay in their response. The last strategy with the above average frequency range was *referencing*. It was inferred to be employed by the subjects for seven out of ten items. They referred to certain words or phrases in the sentence with the deleted word in order to determine the appropriateness of their responses. For example, they noted that the deleted word referred to was "background listening". Since "background listening" was considered as an innate object, the relative pronoun "<u>**what**</u>" was identified as the response for item 7. The word "<u>**what**</u>" then acted as a cataphoric reference for "background listening".

There were two strategies within the **average** frequency range but none within the **below average** frequency range. The two strategies comprised *expression* (f = 20) and *eliminating* (f = 16). Both these strategies were utilised by at least one subject in eight out of ten items. All the four subjects exercised *expression* for item 12 due to the phrasal verb "looking **<u>at</u>**". Only subject C thought that "looking <u>to</u>" was the response because of negative transfer from the native language. Other expressions, such as, "abou<u>t</u> what" (item 5), "how <u>deep</u>", "how <u>much</u>", "how <u>many</u>" (item 8) and "<u>can be</u>"

159

(item 18) were suggested by the subjects even though these were incorrect responses. It could be perceived that the use of this strategy had misled the subjects to think that the responses were part of the correct expressions. Secondly, the subjects utilised *eliminating* strategy to help them select the appropriate response from the options that they had generated themselves. Comments such as, "not right" or "it doesn't sound nice/right", indicated the subjects rejected the options that were not so appropriate based on their own *reasoning* and *auditory monitoring*.

The subjects had applied five strategies within the **low** frequency range. These were *interpreting* (f = 9), *semantics* (f = 8), *substitution* (f = 8), *group retrieval* (f = 6), and *auditory monitoring* (f = 6). It was observed that all the subjects except subject B had utilised *interpreting* in assisting them to solve item 4. Subject A utilised it the most frequently to solve three other items (5, 15, and 18) whereas subject C utilised it to solve one more item, that is, item 7. The subjects attempted to read beyond the lines by using their schemata to enhance their comprehension of the text related to the item. They also sometimes verbalised their thoughts in their native language (Mandarin). At times they had interpreted the hidden message correctly. For instance, subject D understood the context for item 4 to be "what is assumed is different from reality" from the sentence before the deleted word, that is, "I'm assuming your listening is like a blank canvas..." and the sentence after the deleted word "...it is not". Subject A understood this context to be "the first sentence is seldom true". However, subject C misinterpreted this context as "when I talk, you must listen". Hence the utilisation of this strategy might not lead them to the right response. Secondly, *semantics* was observed to be exercised at least once by all four subjects for only four items (3, 15, 18, and 19). This strategy was employed when the subjects were uncertain of the meaning of keywords that would determine the response. These words were noted to be just

160

before or after the deleted word. For example, the word "**analogy**" which appeared after the deleted word (item 3) was assigned meanings such as "example" by subject A; "a suggested form of treatment" by subject B; and "a thing" by subject C. Subject D did not utilise this strategy as she pointed out that an analogy was mentioned earlier in the text and another analogy was being referred to in item 3. Unlike the rest of the subjects she appeared to be familiar with the word "**analogy**" and did not make any attempts to identify its meaning. Thirdly, three subjects, A, B and D, employed *substitution* for four items. They had generated more than one response and were substituting these responses before deciding which response would be most appropriate. For example, subject A had produced three options of her own, namely, "<u>idea</u>", "<u>word</u>" and "<u>words</u>" for item 18. She was observed to be substituting these self-generated options to check the suitability of each option. Fourthly, *group retrieval* was exercised at least once by all four subjects in solving five items. They were observed to have generated words that were related to one another in terms of their word class (e.g. determiners : "<u>a</u>" and "<u>an</u>"; quantity qualifiers : "<u>many</u>" and "<u>much</u>"; conjunctions : "<u>and</u>" and "<u>but</u>"); tense forms (e.g. "<u>is</u>" and "<u>was</u>"); and semantics (e.g. "<u>make</u>" and "<u>cause</u>"; and "<u>but</u>" and "<u>actually</u>"). Lastly, *auditory monitoring* was noted to be employed by only subject D for five items. She had actually realised this strategy when she made comments such as "<u>another</u> sounds more suitable" or "<u>it</u> sounds smoother". This strategy had assisted her in deciding the suitability of her responses. The rest of the three subjects could have utilised this strategy as they read aloud their responses. They might not have made similar comments audibly but their choice of response could be influenced by *auditory monitoring*.

Within the **very low** frequency range, the subjects were inferred to have exploited nine strategies. The first two strategies with a frequency of five were *analysis* and

she had selected the latter which was incorrect responses due to the misapplication of grammatical rules. *Stated* was noted to be utilised by two subjects, A and B, for item 18 only. The suggested responses, "word/s" and "idea" were words that were within the sentence for item 18 or the sentence before it. The last three strategies within the very low frequency range with a frequency of 1 were *focusing*, *linking answers*, and *morphology*. Subject D focused on the task imposed on her when she was aware that just responding to the items was insufficient. She was required to provide an explanation on how she arrived at her response. Subject A linked response for item 6 with item 8 when she repeated the response "subject" (item 6) for item 8. Subject A also exercised *morphology* when she noted that "analogy" was in its singular form.

There were seven strategies that were **absent, non-verbalised, or uninferred**. The subjects were not observed to deduce any grammatical rules from the text given by *induction*. Since all the ten items involved function words, *collocation* could be considered as an inappropriate strategy. This is because words that collocate are normally content words. *Phonology* and *grouping* were not inferred. *Understanding* and *comprehension monitoring* could be exercised without verbalisation as these strategies could occur simultaneously with reading aloud. Lastly, *response monitoring* could be considered non-existent because no options were provided for the subjects to select from.

4.3.3 Research Question 3

Which is the more influential factor in determining the utilisation of strategies : deletion type or task type?

In order to ascertain whether the deletion type or the task type is the more crucial factor in deciding which strategies were utilised by the subjects, it would be necessary to make a

number of comparisons. Firstly, there would be a comparison between the strategies employed to solve content word deletions and function word deletions for cloze with options. Secondly, a similar comparison would be made for cloze without options. Thirdly, the strategies utilised for cloze with options would be compared with cloze without options in solving content words. Fourthly, there would be a similar comparison in solving function words. The strategies would be examined according to the nine categories of frequency range in order to establish likeness and differences between the deletion type or task type. The greater the similarities the less influential the factor would be. In other words, the degree of influence of a certain factor would be proportionate to the degree of variation between the strategies employed in accordance to the nine frequency ranges established.

4.3.3.1 Comparison between strategies utilised for content word items and function word items for cloze with options

Although the task type (cloze with options) was the same for different deletion types (content and function words), there were similarities and variations in the strategies employed by the subjects. The similarities would be discussed first followed by the differences. In addition there will be a comparison between the similarities and differences.

4.3.3.1 (a) Similarities (cloze with options)

With reference to Table 4.7, p.269, it could be noted that there were fifteen similar strategies within six frequency ranges, namely ultra high, high, above average, below average, very low, and absent/ non-verbalised/ uninferred. Firstly, the four strategies (*pre-deletion and post-deletion, testing, substitution,* and *elimination*) within the ultra high frequency range for content word deletions were also observed for the function word deletions. Secondly, *reasoning* and *response monitoring* within the high

frequency range were inferred for both deletion types. These similarities observed could be due to the nature of task type. With the availability of the options, the subjects frequently substituted, and tested each of the options. They also monitored their responses by *elimination* and *reasoning* to ensure they had selected the most appropriate response. Thirdly, the similar strategy present within the above average frequency range for both deletion types was *referencing*. For content word deletions, references were made to other content words in the text that were related to the items. For example, the subjects referred to words, such as, "communication", "conversation", and "reactions" in the text before the items concerned and selected the same words as their responses to these items. Although the strategy was the same, the subjects focused on different aspects of the language where function words were concerned. For instance, they referred to the tense used before item 15. They also referred to the contextual meaning of related words for certain items, such as, "I" and "you" would be converted to "**us**" when used after "which of -" in item 1. Fourthly, *delayed response* was noted within the below average range for both deletion types. With the schemata that only one of the four options would be the correct response, the subjects checked the suitability of each option before deciding upon the response. This was particularly the case when they were unsure of the response. It was observed that this strategy was utilised for eight content word items and only five function word items. This could imply that subjects were more hesitant about their response for content words than function words and hence relied more on the options.

Fifthly, there were four strategies (*induction, word retrieval, auditory monitoring* and *language transfer*) within the very low frequency range that were manifested for both content and function word deletions. The utilisation of *induction* ($f = 4$ for both deletion types) was observed for four content word items. However its utilisation concentrated

on only one item, that is, item5, for function word deletion. The subjects mainly focused on the tense associated with this item. Next *word retrieval* was exercised for one content word item and two function word items. The subjects involved utilised this strategy by retrieving expressions related to the word after the deletion (e.g. "<u>other</u>, <u>one</u>, <u>the</u> days") for content word items. The reverse was true for function word items whereby the subjects concerned retrieved phrasal verbs based on the word before the deletion (e.g. "**available** <u>on</u>" and "**look** <u>with</u>"). They had retrieved these words as alternative responses to the options given. Subjects A and B had utilised this strategy as the initial part of their attempt to solve the items for both deletion types, whereas subject D employed this strategy in the middle of her attempt to solve one content word item. The former two subjects appeared to think of familiar expressions and produce their own responses before checking the options given. If their response was listed as one of the options, it would confirm (to a certain degree) that their response could be the most suitable response. On the other hand, if their response was not among one of the options, they would choose the option which was synonymous to it. However if even this was not possible, they would check their comprehension of the related portion of the text and reconsider the probability of each option as the response. Subject D seemed to utilise this strategy only when she was unsure which one of the four options given was the suitable response. She retrieved a word to form a familiar expression related to the item in order to enhance her comprehension of that particular portion of the text. Then she would try to select one of the options which was most synonymous to the expression she had generated as her response. Although the nature of the task did not require the generation of response through *word retrieval*, three of the subjects still exploited this strategy. Hence, the availability of four options did not deter them from introducing their own responses at a very low frequency range. Next *auditory monitoring* was observed to be employed by three subjects (A, B, and D). When they

could not determine the appropriate response for both deletion types, they exercised this strategy to assist them in their selection of the most suitable response. The last strategy listed within the very low frequency range was *language transfer*. This strategy was utilised by subjects C and D to solve only one item in each type of word. They had generated an expression in Mandarin, which they thought would be appropriate as the response. Then they matched this particular expression with the option that was most synonymous in meaning to the Mandarin expression. They utilised their linguistic knowledge of Mandarin expressions when they were uncertain of the individual meaning and usage of the options given.

Sixthly, the three absent/ non-verbalised/ uninferred strategies for both deletion types were *group retrieval*, *synonym*, and *production modulating*. The reason why the subjects were not observed to have utilised *group retrieval* and *production modulating* could be due to the availability of the options. The subjects did not provide any synonyms using the second language. They preferred to use the non-second languages (Mandarin or the Malay language) to enhance their comprehension or to confirm their choice of response. In other words, they employed the *equivalent* strategy instead.

4.3.3.1 (b) Differences (cloze with options)

The variation in strategy use between deletion types appeared to be differences in terms of strategies employed and their frequency of utilisation. On the whole, the subjects exercised one more strategy for function words (that is, 26 strategies) than content words (that is, 25 strategies). With reference to Table 4.8, p.270, it could be noted that within the ultra high and very high frequencies, the subjects had employed *expression* and *deduction* for function word deletions respectively. However these two strategies were less frequently utilised for content words as they were observed to occur within

the above average frequency range for *expression* and average frequency range for *deduction*. This difference could be due to the fact that the function word deletions were focused on expressions, such as, item 11 (look *at,* **for,** *into,* **around**) and item 14 (care *of, for,* **with,** **about**); semantic markers, such as, item 18 (**instead** of) and other expressions, such as, item 7 (available *by,* **in,** *from, through*). In addition the function word deletions concentrated on the use of pronouns, such as, item 1 (*we,* **us,** *them, they*); relative pronouns, such as, item 12 (*why, when,* **what,** *which*); auxiliaries, such as, item 15 (**is,** *are, was, were*); and determiners, such as, item 20 (**a,** *an, the, that*). All these grammatical items appeared to have created the necessity for the subjects to apply their grammatical knowledge through *deduction.* These two strategies were also exhibited for content word deletions but with lower frequencies. These relatively lower frequencies could be due to fewer occurrences of items related to expressions and grammar. There were six such items for content word deletions whereas there were ten such items for function word deletions.

Secondly, *semantics* was manifested within the very high frequency range for content word deletions but within the low frequency range for function word deletions. This could indicate that the subjects focused more on the meaning for content word deletions. They had to process the meaning of the options before they could select the most appropriate response based on their own comprehension. This would particularly be the case for vocabulary-based options, which have distinct meaning of their own.

Thirdly, there were four strategies utilised for content word deletions that were of one frequency range higher than when they were employed for function word deletions. These strategies were as follows:

❖ *equivalent* (high for content word deletions; above average for function word

deletions)

- ❖ *inferencing* (average for content word deletions; below average for function word deletions)
- ❖ *association* (low for content word deletions; very low for function word deletions)
- ❖ *analysis* (low for content word deletions; very low for function word deletions)

One plausible interpretation of this trait could be that the content word deletions posed a steeper challenge to the subjects in terms of lexical knowledge. Hence the need to exercise these strategies more frequently was manifested.

Fourthly, it was observed that *morphology* was exhibited at average frequency range for content word deletions whereas it was noted to occur at very low frequency range for function word deletions. This variation could be the result of the subjects' effort in gauging the meaning and usage of the options given by examining the morphemes that constituted the options for content word items.

Fifthly, there were six absent, non-verbalised and/or uninferred strategies for content word deletions which were utilised for function word deletions within the low and very low frequency ranges. These strategies were *interpreting, focusing, collocation, syntax, understanding,* and *comprehension monitoring.* The absence of *interpreting* could be that the text related to the content word items was sufficiently supportive towards their comprehension of the meaning conveyed in the text. Conversely, the only one incident of *interpreting* exercised by subject C for item 13 (function word deletion) was for the purpose of enhancing comprehension of the meaning conveyed by the text. Next the subjects concentrated solely on solving the content word items and did not pause to focus on the task requirements which were familiar to them. They were actively trying to obtain the most suitable response. Hence they did not utilise *focusing* for content

word deletions. This was unlike the case when subject A was able to respond automatically to the function word item 15 and then posed the question "But why '**is**'?" She focused on one of the task requirements, that is, to explain how she arrived at her response. The third strategy mentioned was *collocation*. This was absent as the content word deletions did not include any collocations unlike function word deletions which included one allocation, that is, "**instead of**" (item 18). This inclusion had given the opportunity for the subjects to employ *collocation* to solve the item concerned. The subjects appeared to be more meaning focused when solving content word items. Hence they did not seem to have examined the structure of the sentence concerned. On the contrary, there were three accounts where subjects B and D considered the position of the deleted function word in relation to the sentence, such as at the beginning or in the middle of the sentence. This was carried out in order to check if their choice of response was appropriate. The last two strategies, *understanding* and *comprehension monitoring*, would most probably be executed without verbalisation. They could have occurred progressively and simultaneously while the subjects were reading the text aloud repeatedly in order to glean the writer's message or idea. Since the message or idea would normally be imbedded in content words, the subjects' comprehension of the text or their response would not be instantaneous. This was because they would need time to seek the meaning of the message or idea. With function word deletions, the writer's message or idea would be less ambiguous to the subjects. consequently the subjects were able to respond spontaneously to some items; indicating that understanding had actually occurred. When this failed to take place, the subjects employed other strategies. There was one incident when Subject A verbalised the monitoring of her comprehension of a portion of the text related to item 7 by saying that she understood what she had read. *Comprehension monitoring* was not observed for content word deletions. This could be due to the possibility that the subjects were

too intent reading to enhance their comprehension in their attempt to solve the content word items that they did not verbalise the monitoring of their comprehension.

Sixthly, there were five absent, non-verbalised and/or uninferred strategies for function word deletions but were utilised for content word deletions within the low and very low frequency ranges. These strategies were *linking answers, phonology, grouping, alternative response,* and *stated.* This dissimilarity could be due to the inherent nature of the function words deleted which did not generate the necessity to utilise these strategies.

4.3.3.1 (c) Comparison (cloze with options)

In this comparison, the task type was kept constant, that is, only the cloze with options was examined. The two variables compared were content word deletions and function word deletions. If the percentage for similarities was higher than the percentage for variation, the more influential factor would be the constant factor, that is, the task type. In other words, the task type would determine the kind of strategies that would be employed by the subjects. Conversely, if the percentage for variation was higher, the more vital factor would be the deletion type. According to the statistic obtained, the percentage for variation (55.88%) was higher than the percentage for similarities (44.12%). Since there were more differences (11.76% higher than similarities), the more influencing factor in determining the strategies employed by the subjects was the deletion type. In other words, the variation in the deletion type, that is, content or function word deletions, played a more important role in the subjects' selection of strategies when solving cloze with options.

4.3.3.2 Comparison between strategies utilised for content word deletions and function word deletions for cloze without options

The data for this comparison revealed that there were both similarities and differences in the strategies employed by the subjects when they solved the content word items and function word items for cloze without options. The similarities will be explicated before the differences. This will then be followed by a discussion on the comparison between the similarities and variations.

4.3.3.2 (a) <u>Similarities (cloze without options)</u>

With reference to Table 4.9, p.271, the subjects utilised eleven similar strategies within four frequency ranges to solve both deletion types, that is, content word items and function word items. The subjects employed these strategies with the same or approximately the same frequencies even though the deletion types differed. One plausible cause for these similarities could be the intrinsic nature of the task type, which was cloze without options. Since the task type was the same, that is, kept as a constant factor, the subjects exploited similar strategies which were fundamental to solving this particular type of cloze. For example, the five strategies, namely *pre-deletion and post-deletion, testing, production modulating, word retrieval*, and *equivalent* which were observed within the ultra high frequency range provided the evidence to support this assumption. The subjects also utilised *expression* with average frequency for both deletion types. Since no options were provided, the subjects had to depend on familiar expressions to assist them in generating the appropriate responses. Next there were six strategies which were noted within the average, low and very low frequency ranges. The exploitation of *auditory monitoring* within the average frequency range was rather subject specific. This was because it was mainly utilised by subject D to solve five out of ten function word items and four out of ten content word

items. There were also two incidents where subjects A and C employed this strategy in solving content word items 13 and 14 respectively.

The remaining four strategies, namely *analysis*, *morphology*, *focusing*, and *language transfer*, noted within the very low frequency range, were neither subject specific nor item specific. Their sporadic utilisation would most probably indicate that they were exercised as and when they were useful in assisting the subjects to solve the particular items concerned.

Besides these strategies which were observed within the similar frequency range, there were seven strategies which were noted for both deletion types within the absent, non-verbalised and/or uninferred category. These strategies included *induction*, *collocation*, *phonology*, *grouping*, *understanding*, *comprehension monitoring*, and *response monitoring*. There could be four possible interpretations for this occurrence. Firstly, the subjects could have overlooked the strategies, particularly *induction*, since no guidance or cues were provided as indications of what the suitable response could be. In their perplexity to complete the demanding task, it would be acceptable if the subjects had ignored these strategies. Secondly, the items themselves might not generate the need to utilise these strategies, such as *collocation*, *phonology*, and *grouping*. Thirdly, *understanding* and *comprehension monitoring* could have been a silent and continuous experience instead of being verbalised or spontaneous. Fourthly, the integral nature of the task might not require these strategies. As an illustration, *response monitoring* would not be necessary as no options were provided.

4.3.3.2 (b) Differences (cloze without options)

Despite the same task type, several variations in the use of strategies could be noted. With reference to Table 4.10, p.272, there were ten strategies that were utilised more often for content word items than for function word items. The first three strategies, namely *referencing*, *alternative response*, and *inferencing* were observed within the ultra high frequency range for content word items but they were perceived to have been utilised within the above average frequency range for function word items. This observation revealed that the subjects generated more responses for content word items than function word items through *referencing* and *inferencing*. This could be attributed to the fact that content words compose the meaning of the text. The subjects attempted to produce more options from which they could select the most appropriate response to convey the writer's message accurately. Another strategy, *delayed response*, which was within the very high frequency range for content word items was also noted to be within the above average frequency range for function word items. This evidenced that content word items posed a greater challenge to the subjects than function words. The next two strategies, *eliminating* and *interpreting* that were noted within the above average frequency range for content word items. However they were utilised within the average frequency range and low frequency respectively for function word items. Since there were more options generated for content word items, the frequency for eliminating these options would also be higher for content word items than function word items. In order to generate these options, the subjects had to enhance their comprehension of the related portion of the text through *interpreting*. Hence these two strategies were employed more often for content word items than for function word items. The subjects were also observed to have exercised *stated* within the average frequency range, *substitution*, and *association* within the below average frequency range as well as *synonym* within the low frequency range for content word items.

Although these strategies were likewise utilised for function word items, they were noted within the very low frequency range except for *substitution* which existed within the low frequency range. The subjects employed *stated* more frequently for content word items because they referred to certain keywords or important content words in the text and proposed these words as probable responses. The two counts whereby subjects A and B utilised *stated* for function word item 18 were due to their misconception that the deleted word (**it**) was a content word (<u>idea</u>, <u>word</u>). In actual fact, there should be no *stated* employed by the subjects for function word items. Content word items rendered themselves more liberally to *association*, and *synonym* as they were "pockets" of meaning to the text. In other words, the integral meaning of content words and their shades of meaning could trigger *association* and *synonym* more readily than function words. The subjects generated sluggishly more options for content word items than function words. Therefore, the subjects employed *substitution* at a sluggishly higher frequency for content word items than function word items.

Another aspect of variation included four strategies that were noted with higher frequencies for function word items than for content word items. These strategies were *reasoning*, *deduction*, *semantics*, and *group retrieval*. *Reasoning* was observed within the ultra high frequency range for function word items but it was noted within the above average frequency range for content word items. It was a more predominant strategy for function word items due to the subjects' attempt to link the ideas presented by the writer. They had to reason out the writer's message logically prior to generating their own responses. Its lower frequency for content word items could be due to the absence of options which actually functioned as guidance and cues for the appropriate responses. Instead of reasoning the subjects resorted to strategies such as *referencing*, and *inferencing* whereby they could rely on the vocabulary within the text as their

potential responses. Next *deduction* was inferred to be utilised within the above average frequency range while *semantics*, and *group retrieval* were noted within the low frequency range for function word items. However all these three strategies were observed within the very low frequency range for content word items. The subjects applied their grammatical knowledge more frequently for function word items as these items were grammatically oriented. The higher frequency obtained for *semantics* was the outcome of the subjects' attempt to gauge the meaning of the words before and after the item. It was utilised indirectly in assisting them to solve the function word items. The unexpected low frequency of *semantics* for content word items could be attributed to the unavailability of options. It was also for this reason, the subjects employed *group retrieval*. Its more frequent usage for function word items could reveal that it was easier for the subjects to retrieve a group of words (e.g. wh-words, auxiliaries) for these items rather than for content word items.

Last but not least, there were two strategies, namely *linking answers*, and *syntax*, which were observed to be absent, non-verbalised and/or uninferred for content word items. However, these were noted to be utilised within the very low frequency range for function word items. This observation could indicate that the subjects approached function word items at the sentence level, particularly when the deleted function word connected the idea before the item with the idea after it.

4.3.3.2 (c) Comparison (cloze without options)

The percentage for similarities (52.94 %) was calculated to be higher than the percentage for variations (47.06 %). Since the constant factor was task type, it could be concluded that this factor, that is, cloze without options, was the more significant factor

in determining the type of strategies employed by the subjects. Hence the deletion type could be considered as the sluggishly minor factor when no options were given.

4.3.3.3 Comparison between strategies utilised for cloze with options and cloze without options for content words

In this comparison, the deletion type is kept constant. Only the strategies utilised to solve the twenty content word items for both task types (cloze with options and cloze without options) will be examined. The similarities and the differences will be discussed followed by comments on the comparison of these two entities.

4.3.3.3. (a) Similarities (content words)

There were only three frequency ranges whereby seven similar strategies were observed. Table 4.11, p.273 showed that the two most highly utilised strategies within the ultra high frequency range were *pre-deletion and post-deletion*, and *testing*. Even though the task types (cloze with options and cloze without options) differed, these two strategies appeared to be fundamental strategies in solving content word items. *Pre-deletion and post-deletion* was an essential strategy because it enhanced the subjects' comprehension and assisted them in selecting or generating the most suitable response. *Testing* was also frequently employed by the subjects to check the suitability of the option selected or response generated. It was perceived that there were no similar strategies within the next six frequency ranges, that is, from very high frequency range to low frequency range. This phenomenon could be due to the distinctive nature of each task type which placed its specific demands on the subjects, particularly those strategies with a great difference in frequency. The third strategy, *language transfer*, was observed to be utilised within very low frequency range. This strategy appeared to be commonly applied for both task types with similar frequency. This could imply that the

utilisation of *language transfer* was independent of task types. The last four strategies, namely *collocation*, *syntax*, *understanding*, and *comprehension monitoring* were found to be absent, non-verbalised and/or uninferred for both task types. The content words deleted did not render themselves as collocates. Hence, there was no inclination to utilise *collocation*. The subjects were more focused on the words in the text rather than the sentence structure or sentence types as they resolved the items. Thus *syntax* was not observed for both task types. As pointed out earlier, *understanding*, and *comprehension monitoring* could be employed but these strategies would be gradual and they might occur simultaneously with reading aloud for both task types.

4.3.3.3. (b) Differences (content words)

The differences in the frequency of strategies employed to solve content word items for both task types were phenomenal as illustrated by the data in Table 4.12, p.274. There were variations of strategies for all nine frequency ranges. A total of thirteen strategies employed for cloze with options was found to be higher in frequency than when they were employed for cloze without options. consequently, there were nine strategies that were noted to be employed less frequently for cloze with options than cloze without options. In addition there were five strategies that were absent, non-verbalised and/or uninferred for cloze with options but were observed to be employed for cloze without options. These strategies existed within the four different frequency ranges, namely, ultra high (*production modulating*), above average (*interpreting*), low (*synonym*), and very low (*focusing,* and *group retrieval*). Conversely there were also five strategies absent, non-verbalised, and/or uninferred for cloze without options but were noted to be utilised for cloze with options. These included *response monitoring* (high), *grouping* (low), as well as *induction, linking answers ,*and *phonology* (very low).

These variations would be most probably dictated by the task type. The greater the difference in the frequency of the strategies between the two task types, the greater would be the influence of the task type. For example, *production modulating* was absent for cloze with options but it was observed to be employed within the ultra high frequency range by the subjects for cloze without options. Since there was no production required for cloze with options, there would be no production modulating. However, the subjects were required to produce their own response for cloze without options. Therefore they would need to modulate their self-generated responses. It could be perceived that the task requirement for cloze without options pre-eminently differentiated the frequencies of *alternative response, word retrieval, interpreting, inferencing, delayed response, referencing,* and *stated* as being higher than their respective frequencies for cloze with options. Conversely the availability of the options proliferated the use of certain strategies such as *substitution, elimination, semantics, response monitoring,* and *morphology* for the cloze with options. Another cause for the variations could be the inherent nature of the words in the text and the deleted words or options which could encourage the employment of certain strategies, namely, *expression, association, stated, auditory monitoring, linking answers,* and *phonology.*

4.3.3.3. (c) <u>Comparison (content words)</u>

In the comparison of the strategies employed to solve content word items for cloze with options and cloze without options, the variations (79.41 %) outweighed the similarities (20.59 %) by 58.82 %. Hence for content word items, the more pre-eminent factor in establishing strategies to be employed would be the task type and not the deletion type. The frequencies of these strategies were also mainly dictated by the task type.

4.3.3.4 Comparison between strategies utilised for cloze with options and cloze without options for function words

The variables under comparison are the two cloze types, that is, cloze with options and cloze without options. The constant factor for this comparison is function word items. In other words, all the strategies employed to solve the function word items for cloze with options are compared with those utilised for cloze without options. The similarities of this comparison will be examined before the differences. Conclusions drawn from this comparison will also entail.

4.3.3.4 (a) <u>Similarities (function words)</u>

There were more similarities among the strategies utilised for function word items (13 strategies : refer to Table 4.13, p.275) than content word items (7 strategies : Table 4.11, p.273). There were ten strategies inferred or observed to occur with ultra high, above average, low and very low frequencies while three strategies were absent, non-verbalised and/or uninferred.

In other words there were no similar strategies observed within four other frequency ranges, namely, very high, high, average, and below average. The two common strategies within the ultra high frequency range were *pre-deletion* and *post-deletion, and testing*. These two strategies were found to be within this frequency range in all four comparisons for the same reasons. Although the task types differed, the subjects utilised *referencing* with above average frequency to solve function word items for both task types. They attempted to gauge the suitability of their responses by referring to grammatical aspects in the text, such as tenses and pronouns. Moreover the subjects utilised *semantics* (low frequency range), *analysis, syntax, language transfer, association, morphology*, and *focusing* (very low frequency range) for both task types.

The common presence of these strategies at the low frequencies observed could indicate that they would not be specific to the task type. These were utilised as additional means to solve the items due to their particular characteristics. *Phonology, grouping* and *understanding* were the three common strategies that were absent, non-verbalised and/or uninferred. The first two could be deemed unnecessary in solving the items while *understanding* could be nonverbalised.

4.3.3.4 (b) Differences (function words)

The data in Table 4.14, p.276 indicated that there were seven strategies with a minimal difference of four frequency ranges between cloze with options and cloze without options. Within this parameter, it was observed that a total of four strategies were utilised with higher frequencies for cloze with options than cloze without options. The two strategies with the greatest difference in frequency were noted to be *substitution,* and *response monitoring*. *Substitution* was observed to be within the ultra high range for cloze with options while its occurrence was only within the low frequency range for cloze without options. In addition, *response monitoring* was perceived to be manifested within the high frequency range for cloze without options but it was noted as uninferred for cloze without options. On a lesser scale of difference were two more strategies that occurred within the ultra high frequency range for cloze with options but within the average frequency range for cloze without options. These strategies were *expression,* and *elimination*. On the other hand there were also three strategies that were higher in frequency for cloze without options than cloze with options. *Production modulating,* and *word retrieval* were exhibited within the ultra high frequency range for cloze without options. The former strategy was absent while the latter was observed to be utilised within the very low frequency range for cloze with options. The third strategy, *alternative response,* occurred within the above average frequency range for cloze

without options but it was uninferred for cloze with options. These significant dissimilarities would most probably be due to the distinctive characteristics of each task type. For instance, in solving items for cloze with options, the subjects employed *substitution*, *response monitoring*, and *elimination* frequently as they put on trial each option provided. On the other hand, they employed *word retrieval*, *alternative response* and *production modulating* more often in solving items for cloze without options because the absence of options generated the necessity to produce their own responses and modulate their responses.

There were fourteen strategies within a difference of one to three frequency ranges. Only four of these strategies were found to be higher in frequency for cloze with options than cloze without options. They included *deduction, induction, collocation,* and *comprehension monitoring. Deduction* was observed within the very high frequency range for cloze with options but within the above average frequency range for cloze without options. Although the last three strategies were uninferred to be within the very low frequency range for cloze with options, they were non-verbalised or uninferred for cloze without options. Conversely there were ten strategies with higher frequencies for cloze without options than cloze with options. *Equivalent* and *reasoning* were exhibited within the ultra high frequency range for cloze without options. The former was present within the above average frequency range while the latter was found to be within the high frequency range for cloze with options. In addition *inferencing* and *delayed response* which were noted within the above average frequency range for cloze without options were observed to be within the below average frequency range for cloze with options. Furthermore, *interpreting*, and *auditory monitoring* were exercised with low frequency for cloze without options but with very low frequency for cloze with options. Lastly, *group retrieval* within the low frequency

range and *synonym*, *stated* as well as *linking answers* within the very low frequency range for cloze without options were absent, non-verbalised and/or uninferred for cloze with options. The variations in the cloze frequency ranges could be attributed to the intrinsic demands of the individual cloze type. A probable explication for the absence or non-verbalisation of certain strategies could be due to the specific characteristics of the word deleted or the options available which did not initiate these strategies.

4.3.3.4 (c) <u>Comparison (function words)</u>

It was noted that the percentage of similarities among the strategies employed for both task types was 38.24 %. This figure was less than the percentage of variation, which was calculated at 61.76 %. The difference of 23.52 % in favour of variation revealed that the task type was the more significant factor than the deletion type in prescribing the types of strategies utilised by the subjects to solve function word items.

There were four factors that were compared in this study. These four factors could be categorised under two main factors, namely, task type and deletion type. The task type consisted of cloze with options and cloze without options. The deletion type comprised content word deletions and function word deletions. The four comparisons made were aimed at establishing the more dominant factor in determining the types of strategies employed by the four subjects as they solved the two types of cloze. With reference to Table 4.15, p.277 the following conclusions could be drawn:

1. When comparing content word deletions and function word deletions for cloze with options the more significant factor in determining the choice of strategies was the deletion type.

2. When comparing content word deletions and function word deletions for cloze without options, the determining factor in the choice of strategies was the task type.

3. When comparing cloze with options and cloze without options for content words, the more dominant factor in the choice of strategies was the task type.

4. When comparing cloze with options and cloze without options for function words, the more important factor in the choice of strategies was the task type.

5. The two common and most frequently utilized strategies for the four comparisons were *pre-deletion and post-deletion* and *testing*.

The outcome of the comparisons revealed that the more dominant factor in determining the choice of language processing strategies would generally be the task type. Besides the task type and the deletion type, there could be other minor contributing factors that determine the choice of strategies. These might be the words and syntax in the text, the deleted word or options, and the idiosyncrasies of the subjects' preferences in the use of the language processing strategies.

4.3.4 Research Question 4

Do the subjects practise the strategies they have been taught? In other words, is the teaching of cloze completion strategies effective in assisting the subjects in their performance?

Prior to discussing the observations in relation to this research question, the data obtained from interviewing the English language instructors is reviewed.

4.3.4.1 Data obtained from interview with English language instructors

The outcome of the interview sessions with the English language instructors revealed that five out of eight instructors in the Polytechnic of Port Dickson actually taught their students how to complete cloze texts. The apparent rationale for not teaching was the

assumption that the students would know how to complete cloze texts due to their prior experience during their primary and secondary education.

The instructors had used different terms to describe the strategies that they taught their students. They were also very specific in stating the strategies taught. These are summarised in Table 4.16, p.278 (cloze with options) and Table 4.17, p.279 (cloze without options). Since the focus of the study is to investigate how the subjects solve specific items in the cloze texts, preparatory procedures (such as pre-reading activities, and vocabulary exercises) undertaken by the instructors were excluded.

A comparison between the strategies listed in the two tables revealed that there were similarities as well as differences in the strategies taught in solving the two types of cloze. These strategies can be identified as three broad classes, namely, text oriented, language oriented and task oriented strategies. The first two classes were reiterated for both types of cloze while the third differed between the two cloze types. The discussion of the findings will entail two aspects. Firstly, a comparison will be made between the strategies advocated by the instructors and the strategies utilised by the subjects. Secondly, the mismatch between these two sets of strategies will be highlighted with plausible explication.

4.3.4.2 Comparison between strategies advocated by instructors and strategies utilised by subjects

To begin with, the text oriented strategies which were related to reading, comprised contextualisation strategies such as *referencing* and *pre-deletion and post-deletion*. The subjects had employed these strategies in solving the two types of cloze. The frequencies of these strategies are in Table 4.18, p. 280.

In quantitative terms, the frequency for *referencing* indicated that it was utilised pre-eminently (that is 46.77%) for solving content words deleted in the cloze without options. The *pre-deletion and post-deletion* was the most utilised of strategy for both cloze types in the study. The statistics in Table 4.18, p.280 indicated a higher frequency for content words rather than function words. These text oriented strategies could be viewed as being the most pertinent strategies in assisting the subjects to establish an appropriate context before they could respond accordingly.

Next the language oriented strategies included *expression, deduction, semantics, induction, syntax,* and *collocation.* The frequencies of these strategies as employed by the subjects are illustrated in Table 4.19, p.280. There were three language oriented strategies which could be considered as having attained a satisfactory level of utility by the subjects. These are *expression, deduction* and *semantics. Semantics* was employed mainly to solve content words in cloze with options while *expression* and *deduction* were predominantly employed to solve function words in cloze with options. The remaining three strategies with very low frequencies were *collocation, syntax* and *induction. Collocation* recorded the lowest frequency as the words deleted did not render themselves for this strategy except for item 18 (function word, that is, "**instead of**") in the cloze with options. The utilisation of *syntax* was limited to function words for both cloze types where conjunctions were deleted. *Induction* was utilised only in the cloze with options. This seemed quite an unsatisfactory performance as instructors would expect a greater awareness and sensitivity towards this correct response strategy in gauging the response. A plausible reason for this could be the subjects' inability to utilise this strategy due to their insufficient knowledge about the language or lack of confidence in analysing the linguistic aspects such as tenses, concord and cohesiveness of the text. On the surface, these strategies appeared to be applied appropriately.

However, this relevant exploitation of strategies did not necessarily guarantee positive outcomes. To examine the effectiveness of the strategies as utilised by the subjects, the frequencies of correct and wrong responses for each cloze type were tabulated against the four language oriented strategies as in Table 4.20, p.281.

The relatively high percentage of incorrect responses for strategies involving *expression, deduction, semantics,* and *induction,* pointed to the existence of other contributing factors which could determine the correctness of responses besides the exercise of relevant strategies. With reference to the think-aloud protocols, all the four subjects had similar experiences of employing relevant strategies but producing inappropriate responses for both cloze types. It was evident that subjects employed a variety of strategies to respond to an item. Hence, the accuracy of the subjects' responses might not be solely dependent upon the effective use of those strategies listed in Table 4.20, p.281. In other words the main cause for the inappropriate responses might out be the direct result of incompetence in utilising these strategies. It could probably be incompetence in utilising other strategies involved. Examples of these instances would be discussed to reveal probable explanations for inappropriate responses. Firstly, subjects utilised *expression* to solve the deleted half of a phrasal verb, that is, "look _____" (ref: CO/11, that is, cloze with options/ item number 11). Subject D responded by selecting "*at*" because "look *at*" was a common phrase to subject D instead of "look **for**" which was the correct answer. There were also other expressions, such as, "**passive** *way*" (ref: CNO/2/A, that is, cloze without options/ item number 2/ subject A - the correct response was "**passive activity**"), "care *of*" (CO/14/D - "care **about**"), and "*future* days" (CO/19/C - "**alternative days**") which the subjects thought were suitable but were actually inappropriate within the context. A plausible explanation for the subjects to rely on the familiarity of expressions in order

to select their responses could be that they know only a limited range of expressions. Consequently they chose to ignore other expressions which they were uncertain of or they lack knowledge of in terms of meaning. Secondly, examples of errors arising from the employment of *deduction* would unveil further reasons leading to the ineffective application of this particular strategy. For instance, subject D was under the impression that the conjunction "**but**" (CO/13) should not be used to begin a sentence. Subject D's choice to operationalise this "rule" resulted in the elimination of "**but**". Since the correct response was "**but**", the use of *deduction* did not benefit subject D at all. Similarly, subject C misunderstood the function of "*if*" to be contrastive and to be used between positive and negative sentences. Misled by this "rule", subject C's response was inappropriate as the answer was "**but**" (CNO/4). Another example would be subject C's impression that the conjunction "*and*" was used to join two parts of a long sentence. However, "*and*" was not the deleted word but "**is**" (CNO/5). These three instances illustrated that applying the strategies out of context would not be of much avail. It would be particularly so when these "rules" were learnt inadequately. Thirdly, *semantics* had its drawback when the subjects' misconception of the meanings of certain words caused them to arrive at an inappropriate response. For instance, subject D equated the meaning of "**speaking interest**" as "intention" while subject A assigned the meaning of "quietly" to "**without speaking**". Both the responses were inappropriate as the correct response was "**silence**" (CNO/13). Last but not least, *induction* required the subjects to make a simple linguistic analysis of the text in terms of linguistic aspects, such as, tenses, word forms, sentence structures and sentence types. As an illustration, subject A noticed that two of the three words before the blank (CO/5) ended in "ing", namely, "**specialising**" and "**facilitating**". Subject A then decided to follow this trait by selecting "*organizing*". Although subject C made the same observation, she rejected "*organizing*" as the answer. The rational was there were

too many words ending in "ing" in a single sentence and "*organize*" was selected. Both were inappropriate responses as the correct response was "**organizational**". Another instance was when subject C noted the simple past tense form of two verbs ("called" and "asked") after the blank (CO/15) and selected "*was*" as the response. This was inaccurate as the correct response was "**is**". Although *induction* was utilised, it was rendered ineffective due to the subjects' misapplication of the linguistic aspects they had noticed in the text. This problem could be the result of the subjects' failure to consider their responses in totality of the sentence(s) involved and the context of the paragraph or the text as a whole.

The discussion had unveiled several factors that restricted the exploitation of relevant strategies and caused them to malfunction. These factors comprised the subjects' limited exposure to certain language forms (such as phrasal verbs and expressions), overgeneralization of certain grammatical rules, misinterpretation of the linguistic aspects in the text, and errors in associating meaning to keywords. These limitations stemmed from the shortcomings in the subjects' overall proficiency level, particularly in grammar and the lack of reading.

Lastly, the task oriented strategies included *elimination, substitution* and *testing*. Their respective frequency of use by the subjects are tabulated in Table 4.21, p.281. The very high frequencies of the task oriented strategies indicated that they were certainly utilised to a satisfactory extent by the subjects. These strategies, particularly *substitution* and *testing*, were essential in aiding the subjects to gauge and assess the suitability of their responses.

4.3.4.3 The mismatch between strategies advocated by instructors and strategies utilised by subjects

Deduction was advocated by the English language instructors. However the think-aloud protocols of the four subjects disclosed that this strategy was scarcely put into practice. On the other hand, there were strategies not mentioned by the instructors but were actually being executed by the subjects. The discussion that follows will focus on the probable explanations for this mismatch.

The operations of *deductive* taught by instructors included the identification of the type (function/content) and the class (noun, verb, preposition, pronoun, auxiliary, etc.) of the words deleted. These two operations would require a sound linguistic knowledge of the English Language. Since the subjects lacked competency in this area, they were reluctant to exploit these operations for their benefit. This finding resulted from the stimulated recall with the subjects when they were asked if they could identify the options in terms of which word class or part of speech each option belonged to. The subjects showed no confidence at all and admitted their confusion when they were unable to distinguish the word classes that the options belong to. Similarly the subjects often found it difficult to establish which type of word and word class the response should be. Hence, three out of four of the subjects were unwilling to employ these operations despite the obvious advantage that these could assist them in gauging the appropriate responses.

The strategies utilised by the subjects but not stated by the instructors comprised the use of non-second languages (Chinese and Malay languages), *auditory monitoring, association, grouping, interpreting, linking answers,* and *retrieval strategies.* In an attempt to solve the cloze texts, the subjects had to maximise the use of their limited

resources, giving rise to these strategies. There may be differing reasons as to why these strategies had not been listed by the instructors. Driven by the text, language and task oriented strategies, the instructors could have neglected to mention these strategies which might be learner oriented. However it could not be denied that there could be a possibility that these strategies were not taught directly. This in turn could be due to unawareness of these strategies on the instructors' part. For instance, they might find the retrieval strategies too obscure to be taught precisely. They would also normally not encourage the use of non-second languages in solving cloze texts. In addition, the important role of schema in association strategy as well as the intuition of "play by ear" might have been ignored when more tangible strategies were available. Furthermore, to establish links between answers could be considered as a too precarious strategy to teach as one mistake may lead to another. Interpreting the text would be challenging to teach due to the subjectivity of the reader's schema and idiosyncrasies. Despite the fact that these strategies had not been stated by the instructors, they existed in the subjects' realm of language processing strategies and were executed appropriately as and when the need arose.