AN ANALYSIS OF THE LANGUAGE OF THE INTERNATIONAL BACCALAUREATE COMPUTING STUDIES EXAMINATION 1982-1993

by

PAUL SHOEBOTTOM

Submitted in partial fulfilment of the requirements for
the Degree of Master of Arts (TEFL/TESL)

Centre for English Language Studies
University of Birmingham
Birmingham B15 2TT, England

July 1993

ABSTRACT

Many of the private schools which have been established around the world to provide an English-medium education for the children of expatriate families offer the International Baccalaureate (IB) as their matriculation examination. In this paper the language of the IB Computing Studies examination (1982-1993) is analysed to determine its comprehensibility for the many candidates who do not have English as their first language. The analysis is followed by a description of an experiment which was carried out at Frankfurt International School with the intention of testing the hypothesis that adapting questions from the examination would improve their comprehensibi1ity. Although the experiment could not be completed successfully, and consequently the hypothesis could neither be proved nor disproved, the conclusion of the analysis as a whole is that the IB Computing Studies examination is to a significant extent a test of English as well as a test of knowledge about computing.

.....

ACKNOWLEDGMENT

The author acknowledges the permission of the International Baccalaureate to use extracts from the Computing Studies papers 1982-1993. He wishes to make expressly clear that the kinds of problematic language analysed in this study were largely eliminated from the Computing Studies examination papers shortly after this dissertation was finished. The computing studies teacher at Frankfurt International School reports that examinations in recent years have been fair tests of computing knowledge and skills for both native and non-native speakers of English.

.....

NOTE

This study has been reformatted as follows to increase its online readability:

Quotes:	*Times New Roman Italics*
Extracts from the IB papers:	"Courier New"
Suggested reformulations:	Verdana

.....

Paul Shoebottom
December 2005

CONTENTS

1. INTRODUCTION

The steady rise in international trade over the last few decades has resulted in a corresponding increase in the numbers of business people and bankers spending parts of their careers in a foreign country. In many cases they are accompanied by their spouses and children, and stay abroad for periods of two to three years or more before returning home or being relocated to another country.

A very important consideration for expatriate families concerns the best way of continuing their children's schooling. One possibility is to send the children to a local state school, but unless they are proficient in the native language of the country, their education is likely to be severely disrupted. Consequently, most parents seek to enroll their children in a private school where the instruction is in the first language (LI) or in English. This is particularly the case if they know that their stay abroad is likely to be limited.

In some countries with large expatriate communities, from France or Japan for example, there are private schools in which the language of instruction is the language of those communities, but the pre-eminence of English as the medium of international communication, together with the preponderance of American and British expatriates, means that the great majority of such schools around the world have English as the main or only language of instruction.

Although, typically, these international schools are private institutions, run by parent governors and each responsible for its own curriculum, most of the reputable ones are members of either the European Council of International Schools (ECIS) or the Middle States Association of Colleges and Schools (MSACS). These are accreditation bodies set up to monitor standards and ensure a broadly similar program of studies in affiliated schools.

International schools are responsible for deciding which leaving examination should be taken by those students wishing to enter higher education. In some cases the 'A' Level examination or American High School Diploma is offered, particularly by schools with predominantly British or American student populations. However, in a large number of cases (over 400 schools in 60 different countries) the matriculation examination chosen is the International Baccalaureate (IB) - an examination set up over 25 years ago with the express purpose of serving the needs of both native and nonnative English speaking students in international schools. The IB claims to be "a deliberate compromise between the specialization required in some national systems and. the breadth preferred in others", and one of its aims is "to facilitate geographical and cultural mobility". It serves as a qualification for admission to universities throughout the world, so that, for example, a Japanese student at Frankfurt International School can gain entry to a European university on the basis of satisfactory results in his IB examinations.

This research project was prompted by an observation by the computing studies teacher at Frankfurt International School that his nonnative English IB candidates regularly failed to achieve the examination grades they had shown themselves to be capable of during the two years of the study course. It was felt that one of the reasons for this could be that these candidates have disproportionate difficulty in understanding the language of the examination questions. This difficulty, at best, means that they have to spend more time interpreting the exact requirements of the question, or, at worst, that they misinterpret these requirements and fail to answer the question set.

Another possible reason for the postulated relatively poor performance of nonnative candidates is that they may fail to attempt those questions where they have most knowledge and could gain the best marks simply because they do not understand what they have to do, or are not sure that they understand and do not want to take unnecessary risks. O'Donnell (1968), for example, found that the

questions most frequently chosen by the 900 candidates in an '0' Level physics examination were those questions that were least syntactically complex. As Swales (1982) summarizes:

> *All students will stay away from questions containing words or sentences that they are not confident they understand. Thus the overseas student may, if the examiner is not careful, have effectively a more limited choice than his British counterparts.*

The issue of candidates failing to perform to potential because of the difficulty of the language of examination questions rather than because of the difficulty of the question tasks themselves is an important one, and raises serious questions about the validity of that examination. However, in order to substantiate beyond doubt the anecdotal postulation that nonnative entrants to the IB Computing Studies examination do indeed score lower because of their problems with the examination texts a study would be required in which the coursework and examination performance of both native and nonnative candidates in a number of different international schools are monitored and compared over a number of years. Even then the reliability of any conclusions that could be drawn from the study would depend on how carefully the many variables had been taken into account. It would be necessary, for example, to find a way of determining whether candidates who performed poorly did so *because* they failed to understand the questions or *despite* understanding them. To find this out would necessitate some kind of separate comprehension test. It would also be necessary to interview candidates to find out the reasons why they chose the questions they did.

Such a study is beyond the scope of this present project. Instead, the central aim here is to undertake a detailed analysis of the language of the examination texts, based on a sample of the last ten years' papers, and to identify and discuss questions or parts of questions which may be considered difficult for various reasons.

However, before we can discuss the potential difficulty of the language of examination questions, or indeed of any given written text, and the ways it could be simplified, it is important that we have a firm understanding of the process of reading itself. Until we have some idea of how people engage with texts, we will not be in a position to predict when and why comprehension is likely to break down. For this reason, the next section of this paper will contain a review of current theories on how people read. Section 3, which follows it, is devoted to a review of the research into what makes texts difficult, and section 4 consists of a discussion of how they can be made more comprehensible. The detailed analysis of the IB Computing Studies examination questions follows in section 5, and section 6 contains an account of an experiment set up to test the hypothesis that adapting the questions would make them more intelligible. Section 7 discusses the implications for teachers of the analysis in the previous sections, and section 8 concludes the paper.

2. THEORY OF READING

Generally speaking, the current understanding of reading as essentially an interactive process has evolved out of a synthesis of bottom-up and top-down theories. Bottom-up theories, which were dominant until about 20 years ago, viewed reading:

> *.. primarily as a decoding process of reconstructing the author's intended meaning via recognizing the printed letters and words, and building up a meaning for a text from the smallest textual units at the 'bottom' (letters and words) to larger and larger units at the 'top' (phrases, clauses, intersentential linkages).* (Carrel1, 1988a)

Cohen (1990) provides a more detailed analysis of the stages of this bottom-up decoding process. He calls the first stage, in which written symbols are converted into sounds, *recoding*; and he distinguishes between recoding in the first language (LI) and recoding in a second language (L2). When reading in LI, it is usually only necessary to sample certain features of certain letters in a word in order to identify that word; in L2, on the other hand, more sampling is required and the sampling is more susceptible to error, particularly when reading in an L2 with a different orthographical system to LI. Cohen stresses that in this first stage process there is no attention to meaning.

The obtaining of meaning from the words on the page, i.e. decoding, can take place at one or more of four sublevels: grammar, information, discourse and writer's intent. At the grammatical sublevel, decoding takes the form of the identification of words as one of the parts of speech. In LI, of course, this process is most often carried out sub-consciously, and as such differs from the careful, conscious decoding of a difficult text in L2. At the informational sublevel decoding is performed to extract the basic information contained in the text, while discourse level decoding deals with the identification of the cohesive elements in the text, together with the intersentential relationships which give the text coherence. The final sublevel of decoding occurs when the reader engages with the writer's intent, a process which may rely on the identification and interpretation of subtle stylistic devices.

Common to all bottom-up or data driven theories of reading is the view that meaning is reconstructed serially by starting with the visual stimuli, i.e. the letters and words of the printed text, and working upwards. The first researcher to reverse this order and to lay more emphasis on the active role the reader has to play was Goodman (1967) with his much-quoted definition of reading as a *'psycholinguistic guessing game'*. The efficient reader is one whose prior knowledge of syntax and semantics reduces his dependence on the printed stimuli and allows him to decode with minimal text sampling. This sampling enables 'chunks' of meaning to be transferred to the reader's short term memory, which in turn enables him to make sense of the text as his eyes move rapidly across the page. If the sampling does not produce a meaningful chunk, the reader will often read on in the hope that the problem will be resolved in the next section of text. If it is not resolved, and what is not understood seems important, he will be compelled to regress and re-sample the same text extract, this time more carefully, in an effort to get at the meaning. (See Smith, 1978, and Perera, 1982 for good accounts of this process.)

The top-down theories which grew out of these insights share this focus on the reader, and not on the text, as the starting point in the construction of meaning. The theories hold that as we read we are constantly making predictions as to what is to come; sampling of the text takes place only in order to verify or revise our hypotheses and predictions. Smith (1978) summarizes this process as a movement from meanings to words rather than from words to meanings.

An integral part of the top-down model of reading is schema theory, which claims that the predictions we generate about a text, and the way we understand and interpret it, are heavily influenced by our background knowledge of the subject of that text. According to Widdowson (1978), meaning does not reside in the text - a text only has potential meaning, which is realized by the reader according to his reading purpose and prior knowledge. This active role of the reader in decoding the meaning of the text is also emphasized by Clarke and Silberstein (1975):

> *More information is contributed by the reader than by the print on the page. That is, readers understand what they read because they are able to take the stimulus beyond its graphic representation and assign it membership to an appropriate group of concepts already stored in their memories .. The reader brings to the task a formidable amount of information and ideas, attitudes and beliefs. This knowledge coupled with the ability to make linguistic predictions, determines the expectations the reader will develop as he reads.*

The top-down model has certainly been important in redressing some of the problems inherent in the bottom-up theories. For example, a major criticism of a rigid bottom-up view of reading is that a uni-directional upward passing of information from one stage of the decoding process to the next does not allow for the proven facilitating effect of prior knowledge of text content on word recognition and comprehension. (Stanovich, 1980, quoted in Samuels & Kamil, 1988). But the top-down model too is now generally considered inadequate as a comprehensive theory of reading. One of its deficiencies is that it also fails to incorporate important empirical evidence adequately (Weber, 1984). For example, it does not properly account for situations where the reader has no prior knowledge and is thus unable to generate predictions, but is still able to make sense of what he reads. Neither can it accommodate research findings which show that good readers are able to recognize words far more quickly than it takes them to activate the appropriate schema and to make contextual predictions.

Current theories of reading, therefore, tend to stress the interactive nature of the reading process. The term *'interactive'* needs to be treated with caution since it has two somewhat separate meanings (Grabe, 1988). In the first sense it refers to the interaction between the reader and the text:

> *If we describe the construction of meaning as a constant back and forth movement between (i.) some specific knowledge stored in the reader's memory which is brought forth when he is confronted with a specific text and (ii) the text itself, then we arrive at a two-way relation between actualized schematas (sic) and text. On the one hand there is a top-down movement which is knowledge-based and in which the schematas explain the text; on the other hand, there is a bottom-up movement which is text-based and through which the text modifies the schematas.* (Baton & Cornu. 1984)

In the second sense, the interactive nature of the different reading processes or strategies themselves is stressed. In other words, the higher order (top-down) processes interact with and influence lower order (bottom-up) processes, and vice versa, to enable the reader to construct meaning as efficiently as possible (Rumelhart, 1977).

In fact, there is not one but a cluster of interactive models of reading (i.e. interactive in the second sense), none of which can be regarded as definitive or comprehensive. The model which is perhaps of most relevance to L2 reading is Stanovich's interactive-compensatory model, which claims that the reading process is *interactive* in that decoding is based on information provided simultaneously from several knowledge sources, and that it is *compensatory* in that a deficit in any one knowledge source results in heavier reliance on other knowledge sources regardless of their level in the processing hierarchy. In practice this means, for example, that the poor reader who is both inaccurate and slow at word recognition, may be able to compensate for these deficiencies in bottom-up processing by engaging a top-down strategy based on his knowledge of the text topic. Conversely, the reader who knows nothing about the topic but who is skilled in word recognition may be able to rely on bottom-up processes to decode the text, (Samuels & Karoil, 1988).

Before turning to the issue as to what makes a text difficult to understand, we need to focus more directly on the specific nature of reading in a foreign language. In particular, we need to consider in more detail the issue of whether a nonnative's difficulty with an L2 text is a result of his poor reading skills or because of his limited L2 proficiency. In other words, does he have a reading problem or a language problem?

As Alderson (1984) reports, there have been a number of studies which have attempted to determine if LI reading skills are automatically transferred when reading in L2, but these have produced

conflicting results. Jolly (1978), for example, asserts that reading in a foreign language requires the transference of old skills, not the learning of new ones, but Yorio (1971) claims that the reading problems of foreign language readers are due largely to imperfect knowledge of the target language and to native language interference in the reading process. Alderson believes that the learner reading an L2 text faces both a reading problem and language problem, with the language problem being pre-eminent at lower levels of L2 competence. In other words, after reaching a certain threshold level of language proficiency, the nonnative reader is likely to be able to apply his LI reading skills (Cummins, 1984). Conversely, learners' LI reading strategies may be "*short-circuited*" (Clarke, 1988) if their L2 competence is low. Hudson (1988) emphasizes the interrelationship of language proficiency and reading skills:

> *It seems likely that reading skills per se .. affect the degree to which the proficiency ceiling (i.e. language threshold) restricts comprehension, just as the proficiency ceiling may restrict the degree to which good reading skills are applied. In this symbiosis, either component can affect the degree to which the other affects comprehension.*

The theory of a language threshold level is of great relevance to the present study since it is thoroughly reasonable for the IB Examination Board to expect that nonnative candidates will have attained a certain level of proficiency in the English language. Obviously it would be foolish for a school to enter candidates of low English competence and then to complain that they have difficulty understanding the examination texts. However, the expected level of English proficiency is not specified by the IB examiners. This is not considered to be their responsibility and they assume, again quite reasonably, that schools will not enter candidates who are clearly lacking in the necessary language skills.

It would be useful if the threshold theory could provide some guidance on the suitability of candidates, but Cummins, the originator of the theory, neither clearly defines the nature of the linguistic threshold nor suggests the point at which it is reached. Indeed, he states that it probably varies from individual to individual and will be higher or lower depending on the nature of the task he is engaged in (Cummins, 1991), all of which somewhat reduces the usefulness of the notion. A few researchers have tried to quantify a lexical threshold; for example, Laufer (1992) claims that 3,000 words is the minimum vocabulary for a learner to be able to read well in L2 and that no amount of general academic or reading ability can compensate for a smaller vocabulary than this. But there is no agreement on this figure (see Perkins, 1989), and in general there seems to be no great likelihood that the threshold will ever be adequately defined. Indeed, the very existence of a threshold is disputed by some theorists, (see Baker, 1988). Certainly, in our current state of knowledge we are unable to identify what level of language proficiency a nonnative needs in order to be able to compete on an equal basis with native speakers as far as the reading requirements of an examination are concerned.

3. TEXT DIFFICULTY

Let us now consider in more detail what makes texts difficult. It is clear from the discussion about top-down and bottom-up reading processes that any answer must take into account the interrelationship of two crucial variables: the reader and the text. Text difficulty is not an absolute quality that can be considered in isolation from the reader. What is difficult for one reader may be straightforward to another, depending on many variables such as the reader's proficiency in the language of the text, his reading skills, his interest in the topic of the text, his reading purpose, his frame of mind, and so on. However, once we have identified the essential characteristics of the target reader of a particular text and established his reading purpose, we are justified in analysing the intrinsic features of the text itself in order to predict likely difficulties. Generally speaking, a text can cause comprehension problems if it is lexically, syntactically or conceptually difficult, or if its

rhetorical organization is unclear or illogical. Each of these textual features (with the exception of conceptual difficulty) will be discussed separately in the following analysis, but it is important to stress at the outset that it is often their interaction or combination in texts that makes for reading difficulty, rather than any one feature in isolation.

The first feature to be considered is lexis since there seems to be some agreement that vocabulary is the single most important variable contributing to text difficulty. This agreement spans both reading in LI and reading in L2. For example, Harrison (1980) writing about LI text readability states:

> *Ever since the nineteenth century, .. vocabulary has been considered to be the most important factor determining text difficulty .. Research studies consistently find vocabulary to be the surest single predictor of text difficulty.*

Carrell (1988b), writing about L2, states:

> *Correlations between knowledge of word meanings and ability to comprehend passages containing those words are all high and well established ...*

The effects on top-down processes of unknown words in a text are stressed by Cooper (1985) and Johns (1980). Cooper claims that unpracticed readers are severely and disproportionately disadvantaged by poor knowledge of vocabulary, and Johns suggests that if more than approximately 50 words per 1000 words of running text are unfamiliar, the reader's ability to guess correctly from context is likely to be severely reduced. Furthermore, in two studies of direct relevance to this present analysis, Williams & Dallas (1988) established the *'crucial importance'* of vocabulary in ensuring the readability of mathematics, science and social studies textbooks aimed at ESL learners, and O'Donne 11 (1968) found 'fairly conclusive evidence that vocabulary weakness could affect the performance of candidates (in a science examination) on a considerable scale.'

Clearly, then, unknown words in a text are a likely contributory factor in any failure to understand that text completely, but how can a writer predict which words are likely to cause problems for his particular audience? In terms of general vocabulary, Meara (1984) claims that the shorter and more common the word, the more likely it is, under normal circumstances to be known. Indeed the two factors of word length and word frequency are applied as indices of text difficulty in a majority of the most widely used readability formulae. Furthermore, there is evidence that the time taken by nonnative readers in recognizing and decoding infrequent words is considerably higher than that needed by mother-tongue readers. (See Meara, ibid)

An important reason why a weak vocabulary impedes text comprehension is that the reader may have difficulty in identifying the cohesive ties which help to make a text coherent. If, for example, he fails to make the necessary connection between a word and its synonym or hyponym, or if he simply does not know the meaning of a conjunctive word, such as *'whereas'*, or of a discourse marker, such as *'moreover'*, his ability to comprehend the text may be severely impaired. (For a summary of the research in this area, see Carrell, 1988.)

As far as specialized vocabulary is concerned, it appears that technical words do not cause undue problems to nonnative speakers with knowledge in that field. In fact, it is the subtechnical terms which are much more likely to be unknown and hence potentially difficult (Cassels & Johnstone, 1978). Cowan (1974) defines these terms as context-independent words (such as *'function'*, *'relation'*, *'basis'*, *'adequate'*, *'approximately'*, etc.) which occur with high frequency across technical disciplines but which have low frequency in general texts. (See Cohen et al., 1988, for a discussion of why subtechnical vocabulary causes difficulty.)

Turning now to the role played by syntax in the readability of texts, there is evidence that, in general, readers have more difficulty with long and syntactically complex sentences than with shorter, simpler ones. Harrison (1980) explains that complex sentences cause difficulty because they put too great a load on short-term memory and the information processing capacity.

One feature of complex syntax that increases the burden on short-term memory is embedding. Studies have shown that sentences in which the subject is separated from its verb by an embedded clause or phrase, or the verb is separated from its object, can be more difficult to process, (Fodor & Garrett, 1967). Similarly, any deviation from the standard intrasentential structure of subject followed by verb followed by object may make the reader's task more demanding. As Berman (1984) points out:

> *Foreign language readers' comprehension is liable to be impaired by shifts in SVO (subject, verb, object) ordering; for where the typical expectations of the reader .. are violated, his fluency may be disrupted and hence comprehension hindered.*

Berman claims further that nominalization, a typical feature of many specialist English texts, may also increase decoding time and impede understanding, since there is evidence that active forms are easier to comprehend and recall than an abstract noun formed from the verb. So. for example, *'consideration should be given to the problem of . . '* may be less easy to process than *'you should consider the problem of ..'*.

Williams (1984) comments on the frequent use in English for Specific Purposes (ESP) contexts of nominal compounds, such as the following examples taken from the IB Computing Studies examination papers: `'computerized video lending system'`, `'limited computer staff resources'` or `'word processing computer configuration'`. Williams claims that such nominal compounds may cause severe problems for nonnative readers, particularly those whose mother-tongue, such as Spanish for example, does not have this linguistic form.

In a further study of the role of syntax in text readability Cohen et al. (1979) investigated the effects of the presence in a text of *'heavy'* noun phrases. These are phrases which can be as long as 20 or more words but which perform a single grammatical function; the problem for the reader is to perceive such phrases as one unit, since this is a prerequisite to a full understanding of the passage in which they occur. As an example, the first 21 words of the following sentence (up to and including *'computer'*) form one noun phrase representing the subject of the sentence:

```
Particular reference to the data files you have designed and the
forms to be used to transfer data to the computer should be
given.
```

Cohen claims that such phrases are '*predictably problematic*' for nonnative readers.

A typical feature of many heavy noun phrases is ellipsis in the relative clause which qualifies the noun. The ellipsis may consist in the omission of the relative pronoun and of the required form of the auxiliary verb 'to be', (Fodor & Garrett, 1967). As Berman (1984) points out, this stylistic device, while widely used in the English rhetorical tradition, runs counter to the requirements of maximum transparency from the point of view of the foreign language reader. In the following part question from the 1993 IB Computing Studies paper the 15-word noun phrase starting `'the number of days . . '` contains two examples of reduced relative clauses - i.e. qualifying the nouns 'days' and 'concentrations':

> Write an algorithm to count the number of days having pollutant concentrations equal to the maximum concentration during the year.

The extract may be easier to process more quickly if the ellipsed phrase is filled out as follows:

> Write an algorithm to count the number of days *which* have pollutant concentrations *that* are equal to the maximum concentration during the year.

Compressed passages like the first one above, may, Harrison (1980) suggests, have too few clues to allow the reader to reconstruct the intended message quickly and correctly. However, theorists who propound the dominant role of top-down processes in reading are not convinced that syntax, which at most causes difficulties within the sentence, plays a very large role in whole text comprehension (Candlin, 1979, quoted in Cohen et al . , 1979). Herman (1984) agrees that syntax may not be a crucial influence on the ability to understand the main ideas of a passage, but stresses its importance for complete and accurate comprehension of technical texts:

> *If our aim includes students' acquisition of specific information accurately and in detail - which is the case with much scientific or technical material - exact appreciation of syntactic components of each sentence remains an important aim.*

Moving beyond the boundaries of the sentence let us first focus our attention briefly on how the readability of a text is affected by its cohesive features. We have already seen that failure to recognize or understand lexical relationships or connective words and discourse markers can lead to a breakdown in comprehension. Other cohesive features include the use of anaphoric (backward), cataphoric (forward) or exophoric (out of text) reference, substitution and ellipsis (Halliday & Hasan, 1976). Readers who are unable to interpret these devices are likely to have more problems making sense of the text as a whole, (Cohen, 1984). It must be noted however, that there is some disagreement about the importance for comprehension of recognizing cohesive links. Carrell (1982), for example, argues that cohesion is an effect and not a cause of regarding a text as coherent, and claims that top-down strategies including the use of schemata are more important in comprehension than the bottom-up process of identifying cohesive links. Fulcher (1989) after weighing up the evidence suggests that the linguistic (i.e. bottom-up) and psycholinguistic (i.e. top-down) aspects of reading must be viewed as being equally important. It seems reasonable to conclude, therefore, that some texts may indeed be found difficult because they contain ambiguous, insufficient or unidentified cohesive ties.

The other supra-sentential factor influencing text readability is rhetorical organization. L1 research in this area has not yet produced any generalizable insights other than that texts containing propositions organized in a logical sequence are more easily comprehensible than illogically structured texts, (see Harrison, 1980). Urquhart's (1984) experiments with L2 learners showed that the comprehensibility of any given text can be affected by factors other than lexical and syntactic ones. It cannot be said that these are particularly useful insights, however, and they beg the question as to what constitutes the logical organization for any given type of text. To date, there is no grammar which would provide guidance on the most intelligible way to organize the various rhetorical acts of defining, analysing, illustrating, describing etc. (Mackay et al., 1979) It is possible that the growing body of research subsumed under the title of genre analysis will produce some helpful guidelines in the future.

One last crucial factor that may significantly reduce the readability of any particular text is simply that the text is badly written. It may, for example, contain faulty grammar or unintended ambiguity; it may be poorly punctuated and suffer from a lack of clear discourse markers, or indeed from a general lack of coherence; it may be written in a rambling and woolly style with imprecise vocabulary, or, conversely, it may be impossibly terse; it may even contain typographical errors. Poorly written texts will be found difficult by native and nonnative readers alike, but native speakers are probably more likely to identify them as such, and overcome the problems by using their top-down skills. Nonnative speakers, on the other hand, may be lead to assume that their failure to understand is a result of their own deficiencies, and try to compensate by employing a usually counter-productive, word-by-word bottom-up strategy.

4. TEXT SIMPLIFICATION

If a text is judged as difficult for a particular reader or group of readers, it can be simplified with a view to making it more comprehensible, but this begs the important question: What do we mean by 'comprehensible'? Clearly, we cannot claim to have made a text more comprehensible unless we have an adequate definition of comprehensibility. If we consult the dictionary in an attempt to answer the question, we find 'comprehensible' defined as 'easily understood' (Sinclair, 1987). This is a useful starting point, but we need to clarify further what is meant by 'easily'. It might be 'easy', for example, to decode a foreign language text containing many unknown idioms by using a glossary, but we would probably not want to claim that that text had high comprehensibility. In fact, 'quickly' appears to be a more serviceable term in this context than 'easily', since no text can be regarded as easy if it cannot be decoded rapidly. So a comprehensible text may be regarded as one which can be understood quickly.

As far as the second component of 'comprehensible' is concerned, namely the term 'understood', it is much more difficult to find an adequate definition. The question - *What does it mean to say we have understood a text?* - is unanswerable in general terms. As Widdowson (1990) points out: "*Comprehension is never complete: it is always only approximate, and relative to purpose*." So someone, for example, who reads a newspaper article simply in order to gain certain factual details about a new government policy - which he does to his satisfaction - will claim to have understood that article. But the journalist, whose polemic intent has been overlooked, may disagree.

Since there is no generalizable definition of understanding, we need to define it context by context. Fortunately, this is not difficult in an examination context, where presumably the writer's and the reader's intentions coincide and where understanding can be defined as the correct interpretation of the task to be performed. In other words, a comprehensible examination question can be defined as one which allows the candidate to understand quickly and accurately what he has to do.

Having briefly discussed the object or ends of text simplification, namely improved comprehensibility, we are now in a better position to consider the means by which this object can be achieved. It should be noted at the outset, however, that many specialists in the field (Klare 1974, Honey field, 1977, Davies, 1985, and Widdowson 1978, for example) are sceptical of the effectiveness of text simplification especially when it is carried out by reference to a readability formula and merely entails reducing word and sentence length. They claim that at best such changes have no effect on comprehensibility, and at worst they may even reduce it. As Swaffar (1991) points out:

> *Simplification often alters textual patterns of redundancy and rhetorical organisation, thereby rendering the edited passage less readable than the original*

As far as simplification of syntax is concerned, for example, Blau (1982) found that eighth-grade Puerto Rican students of English had more difficulty with simplified passages consisting of short sentences than with the original version containing complex structures. Her explanation of this is "*that choppy, unnatural sentences are difficult to read, as the relationships and meaning revealed by the formation of complex sentences are probably lost*." On the other hand, there is a large body of evidence that simplifying syntactically difficult texts can and does make them easier to comprehend. (For example, see Brown, 1987, Johnson, 1987, Perera, 1982, and Evans, 1988.)

In order to resolve this apparent paradox, we need to point out that studies into the effects of simplification are necessarily based on a particular group of readers with a particular reading purpose. Therefore it is not surprising if there is no uniformity in the findings. Ulijn & Strother (1990) make this point in their discussion of a study involving university students. While denying the usefulness of syntactic simplification of technical texts at advanced levels, they concede that there is strong evidence that young readers of general texts do find short, simplified sentences easier to understand than longer, more complex ones.

Another reason for the divergent views about the efficacy of simplification lies in the inherent ambiguity of the word *'simplification'*. Simplification can have the text as its focus and be achieved by means of reducing that text to a string of simple (i.e. not compound or complex) sentences and replacing long words with short ones. In this type of simplification it is not difficult to see how the reader's problems may, in certain circumstances, be increased rather than decreased. But the term *'simplification'* can also be used as a synonym for *'adaptation'* and have the reader as its focus - in other words, a text is adapted so that it becomes simpler for the reader to understand, but it does not necessarily become lexically or syntactically simpler. Long et al. (1991) have coined the term *'elaboration'* for this second type of text simplification, which, as they point out, has the considerable advantage of being *'natural'* language in the way that short and choppy simplified texts are not.

Clearly, then, the principles by which texts are simplified must be established context by context. As far as lexis is concerned, for example, the decision about which words to select for simplification depends to a great extent on the reader's background knowledge of the subject, and on his reading purpose (see Alderson & Urquhart, 1964). It also depends on the number of words in the text that are predicted to be unknown. It is clear that the principles of selection for simplification of a specialist text to be scanned for general information by a homogeneous reading group knowledgeable in that field will be quite different from the principles to be applied in the simplification of a general text to be read for detailed comprehension by a heterogeneous group. In the case of an examination, where the texts must be fully understood in the shortest possible time, the most appropriate principle would seem to be to replace less frequent words by their more frequent equivalents, if such equivalents exist. This is particularly so in an examination such as the International Baccalaureate where the candidates are of mixed nationality and English proficiency. There is justification for simplifying lexis in this way in the results of an investigation by Cassels and Johnstone (1978) involving 6,000 secondary level pupils taking a chemistry test. Cassels and Johnstone found that test scores could be *'greatly improved'* by minor, often one-word changes in lexis. They conclude, significantly for this project:

> *It was clear that if pupils failed to answer a question correctly, we could not assume that only their chemistry was faulty. .. It may be that some pupils are prevented from exhibiting their chemical knowledge because the language of the question blocks the process.*

We have seen above that the syntactic simplification of texts is a controversial area, and as with lexis, any decision as to what needs adapting, and how, must take into account all the variables of the particular reading situation. The principle adopted, in the present context is that syntactic simplification can improve the comprehensibility of the examination texts under study, in the sense

of making them easier to decode quickly and accurately. It can do this in three main ways, none of which necessarily entails reducing sentence length: namely, by replacing nominalizations with verb phrases, by relocating embedded phrases (particularly those that separate subject from verb or verb from object), and by expanding reduced relative clauses to include the relative pronoun and auxiliary verb, (see Beaumont, 1982, for a justification of this). So, for example, in the following three extracts from the IB Computing Studies examination, it is considered that:

> On arrival at the check-out point, the customer gives his registered number. - (nominalization)
>
> You may assume an existing subalgorithm, RAISEALARM, which when called starts an alarm bell ringing. – (embedding)
>
> In each case clearly indicate the identifiers needed to select the desired records. - (reduced relative clause)

may be less easy to process, respectively, than:

> When the customer arrives at the check-out, he gives (states) his registered number.
>
> You may assume an existing subalgorithm, RAISEALARM, which starts an alarm bell ringing when it is called.
>
> In each case clearly indicate the identifiers that are needed to select the desired records.

In the first adapted extract above, the topic of the sentence (i.e. 'the customer') has been fronted. Fronting of the subject is another type of syntactic adaptation that can often simplify the reader's task by reducing the load on his short-term memory.

As far as cohesion is concerned, one of the principles by which we can simplify texts is by direct lexical repetition instead of using a synonym. This may make the texts stylistically less agreeable, but in the case of an examination style is not the major consideration. In addition, texts can be adapted to ensure that the link between pronouns and their referents can be made quickly and unambiguously; ellipsis can be 'filled out', as in the second example above where the reduced clause 'when called' is expanded to its full form 'when it is called'. And finally, the organisation of the text and relationships between sentences can be made explicit by the inclusion or highlighting of conjunctions and other discourse markers.

The preceding discussion has focused on the principles by which texts can be simplified - particularly examination texts that are perceived as being too difficult for a particular reader or group of readers. The clear assumption has been that the texts themselves are well written, or at least not badly written, and would be understood without significant problems by those readers for whom they were originally intended. However, as noted earlier, texts can also be difficult because they are poorly written. To make such texts more intelligible to the reader is more a matter of amelioration than of simplification, but since the symptoms of bad writing are manifold, it is not possible to generalize how this can be done. This must be decided context by context, and could involve correcting faulty punctuation or grammar, improving coherence, removing ambiguity, and so on.

Although difficult lexis, complex syntax, and thin cohesion, together with the features of bad writing, have been identified as the most suitable objects of text adaptation, this is not to imply that any of these in isolation will inevitably cause problems, or that simplification will necessarily remove the problems. Indeed, as Parker and Chaudron (1987) state, single adjustments are unlikely to affect the comprehensibi1ity of whole passages. The reader is more likely to experience reading problems or even a complete breakdown in comprehension when a text has an accumulation of the various difficulty factors, to such an extent that the engagement of top-down skills is not enough to compensate for the failure of bottom-up strategies, or vice versa.

Nevertheless, it may well be that in cases where full and detailed comprehension is required quickly and reliably, as in an examination, even isolated difficulties will interfere with accurate and fluent reading. Every such local difficulty may at best distract the candidate and increase the amount of time he needs to work out what he has to do. At worst it may result in him failing to answer the question set, or in giving up on that question and choosing another one about which he may not be so knowledgeable.

For this reason, the following examination of the language of the IB Computing Studies question texts will start with an analysis of a number of short extracts, each one of which contains an example of the types of problem discussed above, including the difficulties specifically associated with poor writing. In each case the original extract will be followed by a suggestion as to how it could be revised to make it more comprehensible, but without in any way changing the nature of the task the candidate has to perform.

5. ANALYSIS OF THE IB COMPUTING STUDIES EXAMINATION PAPERS

5.1 Lexis

The difficulty in the IB Computing Studies papers is not with the specialized lexis; indeed, an integral part of learning a subject entails learning its particular language, (Johns & Dudley-Evans, 1988), and candidates who do not know its specialized terms are rightly penalized in examinations. Rather, the difficulty lies in the use of more formal or subtechnical vocabulary which is not normally encountered in colloquial English.

The underlined words in the following extracts are examples of this type of less frequent lexis. In each case the more common equivalent is shown at the end of the sentence, followed in some cases by a brief discussion. (The numbers in brackets indicate the examination paper containing the extract in question - see Appendix 2)

"Describe briefly three distinct types of input device .." (1) - *different*

The problem with 'distinct' is that it has more than one meaning. 'Different' is preferable, because it is the more frequent word and it is unambiguous.

.....

".. with a fifty per cent growth in the number of retail outlets .." (1) – *(retail) shops*

In the first section of the question in which this extract is situated the reference is to 'retail shops'. The switch to the less common expression 'retail outlets' is not only

unnecessary but may also mislead, the candidate into thinking that some other kind of retail organization is being referred to.

.....

" .. the program will print out the totals of each currency purchased and sold." (2) - *bought*

As wel1 as being a less frequent word than 'bought', 'purchased' is also collocationally odd in conjunction with "sold" .

.....

"A and F are window seats whilst C and D are adjacent to the aisle of the plane." (2) *while / next*

Some candidates may be unfamiliar with this conjunctive meaning of 'whilst'. In fact, there is no need at all for it in this context; a semi-colon or 'and' preceded by a comma would suffice.

.....

In the next example 'whilst' has a different, and perhaps more familiar, meaning. Nevertheless, it has a slightly archaic feel that is not so appropriate in an examination and can be replaced by the more frequent 'while'. Similar comments apply to the two further examples which follow:

"Failures occur during compilation and whilst the program is being run." (3) - *while*

"Amongst the files held on the computer, there is to be: " (6) - *among*

"Thus, 73549 might be keyed in as 75349." (4) - *so*

.....

"The management propose (sic) to install an addition to the system, whereby the guests may use a numerical keypad C to access the central computer. They may also request a final account prior to leaving." (5) - *plans / before*

To avoid using "whereby", the passage could be rewritten as:

The management plans to install a numerical keypad to access the central computer. Guests may use the keypad to ask for a final account before they leave.

.....

"A program accesses a substantial file on disk." (12) - *large*

"The program finishes by producing a printed list of all amendments made to records .." (12) – *changes*

.....

"You may find the list of variable names below and their descriptions helpful in formulating your answer." (14) - *writing*

A non-native candidate may interpret 'formulating' as meaning that his answer has to be given in a particular form. In such a case he may waste time looking up a word which can replaced by an unambiguous alternative.

.....

"Provide a trace of all calls of your subalgorithm which are a consequence of the inital (sic) call from the main algorithm." (20) – *result / first*

The spelling mistake is unfortunate and may lead the candidate to waste time looking up a non-existent word - in any case 'first' would be easier here.

.....

" .. keep the original relational operators intact." (22)

The possible difficulty of the instruction 'keep .. intact' is compounded by the distance of the two words from each other. Easier is:

.. do not change the original relational operators.

.....

"Charles Babbage is usually regarded as the father of the modern computer. The basic principles he deemed necessary in his 19th century designs of the Difference Engine and Analytical engine still define today what a computer is, or is not." (23)

'Deemed' is an uncommon word which may not be known by many candidates. It could be replaced by 'thought', or better still omitted altogether, since it is superfluous in the above context. The revised version reads:

Charles Babbage is usually regarded as the father of the modem computer. The basic principles of his 19th century designs of the Difference Engine and Analytical Engine still define today what a computer is, or is not.

.....

The following extract is an example of a similar problem, 'hence' being both uncommon and unnecessary:

"COPY and complete the table below and hence, by considering all possibilities, show that ... " (25)

.....

"COPY and complete the tables on the facing page indicating the contents of each variable.." (25)

The meaning of 'facing' may not be immediately obvious to some nonnative candidates. In fact, there seems no reason why the actual page number should not be given. Furthermore, the lack of a comma after 'page' allows the interpretation that 'indicating the contents ..' is a relative clause referring to 'page' rather than an amplifying instruction referring to 'copy'. The extract is better rewritten as follows, removing the superfluous capitalization:

Copy and complete the tables on page 5. State the contents of each variable..

.....

"Give at least two advantages of having a hard disk as opposed

to a floppy disk." (15)

'As opposed to' is a difficult expression and a less frequent collocate of 'advantage' than 'over'. However, since the instruction is in itself potentially ambiguous, the extract could be reformulated;

Give two reasons why it is better to have a hard disk than a floppy disk.

.....

The following instruction is very similar:

"Give three advantages of a high level language vs. a low level language." (15)

This time it is the abbreviation that may cause difficulty; the extract can be rephrased in a similar way:

State three ways that a high level language is better than a low level language..

.....

It is not suggested that all or any of the underlined words in the extracts above would cause the nonnative candidate insurmountable difficulty in decoding the meaning of the passage in which occur. The problem lies in the *unnecessary* increase in the processing time required. As Swaffar (1991) points out;

> *When vocabulary meaning is instantaneously recognized (i.e., processed automatically without the conscious effort of the reader), comprehension is faster. The less time spent on .. individual words, the more memory is available to attend to other meaning factors.*

If the candidate does not know a word, however, and is not sufficiently confident that he can guess it correctly from the context, he has to make the decision as to whether it is important or not. If he decides that it is important, then he must look it up in his dictionary - the use of a *'simple translating dictionary'* is allowed in the IB Computing Studies examination. However, consulting a dictionary, even if it can be done accurately, wastes even more precious time. The problem is avoided if the more common word is used in the first place.

5.2 Nominalization

Turning our attention now to syntax, we shall start with an analysis of the difficulties caused by nominalization. Nominalization is a typical feature of formal written language which is appropriate

in many contexts but not necessarily in examinations, where considerations of clarity should always take precedence over considerations of style. The following extracts contain examples of nominalization which may be difficult to understand, and which in most cases can be simplified by their replacement with a phrase containing an active verb.

"Particular reference to the data files you have designed and the forms to be used to transfer data to the computer should be given." (1)

The desire to use a passive construction here has produced the nominalization and resulted in a sentence which is difficult to untangle, particularly because of the distance between the head noun of the subject (reference) and its predicate (should be given). In any case it is not clear why the personal pronoun needs to be avoided since the first sentence of the question starts:

"You have been employed to assist in the preparation .."

Furthermore, candidates may be confused by the meaning of the qualifier 'particular' in 'particular reference'. As this word adds nothing of importance to the instruction, it should be omitted. A more comprehensible version reads as follows:

You should refer a.) to the data files you have designed, and b.) to the forms to be used to transfer data to the computer.

.....

"Gates control entry to and exit from each station. ... Tickets may only be used on the day of issue. On entry, the fact of use of the ticket is recorded thereon." (6)

The last sentence of this extract is difficult and unnatural because of the cumbersome noun phrase 'the fact of use of the ticket' and the formal adverb 'thereon', which is unlikely to be known by many nonnative candidates. A more intelligible, slightly expanded version is:

Gates control entry to and exit from each station. ... Tickets may be used only on the day they are bought. When the traveller goes through the gate, his ticket is marked to show that it has been used.

('Traveller' is the word used elsewhere in the original question, although 'passenger' would appear to be a more natural term for someone travelling by underground.)

.....

"Suggest four instructions which the bank should give the customer concerning the safekeeping of the identification card." (10)

The difficulty of the nominalization is probably compounded by its distance from 'instructions', the noun it qualifies. The following adapted version is likely to be easier to decode quickly:

Suggest four instructions which the bank should give the customer about how to keep the identification card safe.

.....

"The examination board is considering the issue of results to schools by sending back the entry disc .." (19)

It is possible that candidates will initially read the words 'is considering the issue of . . ' as one chunk meaning *'thinking about the question or matter of* ..', whereas in fact 'issue' means 'dissemination' in this context. This potential momentary confusion can be avoided if the extract is rewritten:

The examination board is thinking about issuing the results to schools by sending back the entry disc...

.....

Assume the employee Jane Hausmann marries John Smith and changes her name and address. Upon receipt of the changes, the Payroll Department process them and forward the changes to the Personnel Department, who in turn complete the changes. The resulting record occurrences, after these changes are processed are given below." (20)

The first nominal group 'upon receipt of the changes' is potentially difficult, particularly as some candidates may have encountered the word 'receipt' only in its more common sense of a paper record of payments made. The second nominal group, 'the resulting record occurrences', is opaque and tautologous. In fact, the whole passage needs revision as there is also a problem in the faulty agreement of the subjects 'Payroll Department' and 'Personnel Department' with the plural verbs 'process/forward' and 'complete'. Candidates may also be confused by the relative pronoun 'who' instead of 'which' in reference to the Personnel Department, and by the missing comma in the last sentence after 'processed'. There is also a great suspicion that much of the information in the extract is unnecessary. The following is a severely excised version:

Assume that employee Jane Hausmann marries John Smith and changes her

name and address. These changes are processed by both the Payroll and the Personnel Departments to give the following new records:

.....

The effect of the above changes has been to render highly formal language more colloquial and hence more immediately comprehensible to nonnative candidates. The objection that a colloquial style is not appropriate for an examination is invalid unless the intention is to test question comprehension as well as subject knowledge.

5.3 Embedding

Another feature of syntax that is potentially difficult is embedding. In the following extracts some of the 'embedded' strings (shown by underlining) are marked off by commas, which helps processing. However, each extract can be made readily comprehensible if the word order is changed to avoid the separation of subject and verb, or of verb and object .

"Identify from the above list files which must be accessed serially and files which must be capable of random access."

The embedding of 'from the above list' may cause the candidate to read 'list' and 'files' as one chunk. Furthermore, the instruction word 'identify' is superfluous and the task can be expressed interrogatively:

Which files in the list above must be accessed serially, and which files must be capable of random access?

.....

"Discuss, from the point of view of cost and method of processing, the suitability of magnetic tape compared with disk storage." (5)

The embedded string separates the verb from its object by a long way and makes the instruction less immediately comprehensible. The question is probably easier to understand quickly if the tasks are itemized and the syntax is simplified:

Compare storage on magnetic tape and storage on disk from the point of view of: i.) cost ii.) method of processing.

.....

"Using the programming language you have named in (a), write a single section of program that, given a string variable containing a valid date in either format 2 or format 3, will assign to a second string variable the same date in format 1." (14)

This instruction is difficult to interpret for two reasons. Firstly, the embedded clause starting 'given . . ' separates the head noun (program) by a long way from its verb (assign). And secondly, the embedding of the indirect object, 'to a second string variable' between the verb and its direct object disrupts the natural word order in such sentences. In other words, *'He gave a book to his son'*, is more natural ordering of the direct and indirect objects in a sentence than 'He *gave to his son a book.* ' (See section 5.7 for a more detailed discussion of naturalness.)

A re-ordering and filling out of the extract should give a more comprehensible text:

Assume a string variable contains a valid date in either format 2 or format 3. Using the programming language you have named in (a), write a single section of program that will change this date to format 1 and assign it to a second string variable.

.....

"Explain using a clearly labelled diagram the concept of a linear linked list." (14)

Here again the intrusion of the underlined clause between the verb and its direct object make this instruction less comprehensible than it might be if expressed in two sentences as follows, removing the superfluous 'clearly':

What is a linear linked list? Draw a labelled diagram to explain your answer.

.....

"Sketch a graph (list length in divisions of 15 for the x-axis; time in microseconds for the y-axis) to illustrate relative efficiency for these two methods applied to lists of longer lengths, 15, 30, 45, 60, 75, 90.....180." (24)

The words in brackets are an amplifying instruction which interrupts the main task, 'Sketch a graph to illustrate ..'. Therefore they are better placed at the end of the passage. There are, however, some further difficulties which could be removed. The first concerns the two examples of a kind of shorthand language:

1. the first word in the brackets, 'list', is not preceded by an article and could initially be

misinterpreted as a verb - 'list' is often used as an instruction verb in these IB papers;

2. there is no article in front of 'relative efficiency';

The second difficulty is the complex noun phrase 'these two methods applied to lists of longer lengths', which could be made more immediately transparent by expanding it slightly to include the conjunction 'when' plus a pronoun and an auxiliary verb.

And finally, 'sketch a graph' implies that nothing too rigorous is required as an answer, whereas, in fact, the coordinates have to be plotted exactly. The more precise term needed here is 'draw'.

The following text is the result of these suggested amendments:

Draw a graph to illustrate the relative efficiency of these two methods when they are applied to lists of longer lengths. (The x-axis should show the length of the lists in divisions of 15, from 0 up to 180. The y-axis should show the time in microseconds).

.....

Finally, two examples in which the embedded clause is not properly punctuated:

"Design in flowchart form or using pseudocode, an algorithm to carry out this simulation." (14)

"Explain clearly, using at least three different examples why this statement is true." (25)

In any case, the instructions may be more immediately intelligible if the embedded amplifying instructions are relocated at the end as separate sentences:

Design an algorithm to carry out this simulation. The algorithm should be written in pseudocode or flowchart form.

Explain why this statement is true. Give three different examples.

5.4 Reference

In the following extracts we shall see how candidates may have problems if the cohesive features of texts cannot be identified or comprehended. The focus will be almost exclusively on the difficulties

that may be caused if the reader is unable to make a quick and accurate connection between pronouns and their antecedents, either because there is more than one possible antecedent or because the reference is inexplicit.

"Some companies keep records of customers who repay their loans promptly as being credit-worthy. They make this information available to other companies who make similar loans. Those who are not listed as credit-worthy may find difficulty in obtaining a loan." (8)

'Those' is separated by a sentence from its referent 'customers'. This makes comprehension of the passage unnecessarily difficult, particularly as 'some companies' and 'other companies' are other potential referents of 'those'. The passage can be made more immediately intelligible by replacing the pronoun with the noun it stands for. The revised text, with some further minor amendments in the last sentence, reads:

Some companies keep records of customers who repay their loans on time as being credit-worthy. They make this information available to other companies who make similar loans. Customers who are not listed as credit-worthy may find it difficult to get another loan.

.....

"During the milking process a feeding trough is filled under computer control with an amount of concentrated food directly proportional to the weight of the cow. This is supplemented by powdered additives the quantity of which is directly proportional to the amount of milk produced by the cow over the previous three days." (12)

This is a demanding text, both linguistically and conceptually; linguistically because there is minimal punctuation and the lexis is difficult, and conceptually because the candidate has to juggle with two separate proportional quantities. The intelligibility of the passage is not improved by the inexplicitness of the pronoun 'this'. If the candidate is unsure of the meaning of 'powdered additives', it is very difficult for him to work out what is being referred to, since 'milking process', 'feeding trough', 'computer control', 'concentrated food' or 'cow' are all potential referents. This ambiguity could be avoided by the simple expedient of repeating the word 'food'. The revised text below shows further changes which are considered necessary to render the passage more easily comprehensible. For example, the empty noun 'process' has been removed and the clause it is in has been rewritten; the difficult relative clause starting 'the quality of which..' has been reformulated; and the missing relative pronouns have been restored:

While the cow is being milked, its feeding trough is filled, under computer control, with an amount of food which is directly proportional to the cow's weight. This food is supplemented by powdered additives. The quantity of these additives is directly proportional to the amount of milk that the cow

has produced in the previous three days.

The second sentence of the revised text, starting 'This food ..', is still difficult, but can only be made more intelligible by making fairly radical alterations, including changing part of the content. The second revised version below may not reflect the precise reality of how cows are fed, but the candidate will probably have less difficulty understanding the exact background to the task he has to perform:

While the cow is being milked, its feeding trough is filled, under computer control, with an amount of food which is directly proportional to the cow's weight. Some vitamins are then added to the feeding trough. The quantity of these vitamins is directly proportional to the amount of milk that the cow has produced in the previous three days.

.....

"If a value is not within the inclusive bounds stated earlier, it is considered out of bounds." (18)

Instead of the vague anaphoric reference in the underlined phrase, it would be more helpful to candidates if the bounds were repeated, especially as the amount of text space taken up would be less. Candidates would then not have to search back in the previous text to find the information they need:

If a value is not within the inclusive bounds -99 to +99, it is considered out of bounds.

.....

"Given the flowchart below, redraw this flowchart fragment using the standard flowchart constructs for while/do .." (18)

The following comments about this question were made in the examiner's report:

I feel that the candidates do not read the question thoroughly but read enough of the question to begin a response and do not go back to see if their solution is responsive to the original question. This was very evident in part 2(c) where the candidates tried to redraw the wrong thing. 2(c) clearly states: "Given the flowchart below, redraw this flowchart fragment .." Candidates attempted to draw the pseudocode and not the flowchart!

International Baccalaureate Chief Examiner's Report - May 1991

This is a good example of how ambiguity of reference can result in candidates totally failing to understand what they have to do. It is simply not clear that 'the flowchart below' and 'this

flowchart fragment' are one and the same thing. The confusion is due in part to the inclusion of the first phrase, which is superfluous, and the resulting repetition of the word 'flowchart'. The confusion is compounded by the fact that the flowchart shown later in the question is not identified explicitly as a fragment, whereas the pseudocode shown earlier is. So the candidate can hardly be blamed for misinterpreting the cataphoric (forward) reference implicit in the word 'this' as an anaphoric (backward) one.

The following is a more lucid version of the instruction and contains an explicit reference to the location of the flowchart to be redrawn:

Redraw the flowchart fragment in Figure 1 below. Use the standard flowchart constructs *for* and *while/do* ..

.....

"A bar code is often found on produce sold in supermarkets and, by means of a bar code reader, a computer can directly identify that item. A bar code consists of a series of light and dark lines of varying widths." (18)

The difficulty here is caused by the faulty determiner 'that'. The problem is that 'produce' is a mass noun referring to the total of food on sale in a supermarket and 'item' is a count noun referring to one entity of that total. A nonnative candidate may not immediately make the required connection between the two words. The potential ambiguity is avoided if the passage is reformulated in a more logical order as follows;

Many supermarkets have computers which use a bar code reader to identify food products. The bar code, which is printed on the packaging of most products, consists of a series of light and dark lines of varying widths.

.....

"The scripts are sent straight to the examiners by the school after the examination. After the scripts have been marked the examiner writes the marks down on a form and sends it back to the examining board where they are typed into the computer system. This is a slow and tedious method." (19)

Because of the distance between 'they' and its referent (marks), and the presence between them of another pronoun, 'it', which refers to the form, this passage is not immediately intelligible. The candidate who predicts that the sentence will unfold in one of the two following ways:

After the scripts have been marked the examiner writes the marks down on a form and sends them back to the examining board where

they are typed into the computer system,

or:

After the scripts have been marked the examiner writes the marks down on a form and sends it back to the examining board where it is typed into the computer system,

will have to backtrack in order to reprocess the meaning of 'it' or 'they' respectively.

Furthermore, the difficulty of the extract is possibly compounded for some candidates by the fact that the distinction between the examination board and the examiners is nowhere made explicit - the candidate without the necessary background knowledge (schema) may have problems in understanding why the examiners have to send the scripts to the examining board.

The text is probably more comprehensible as follows. As well as reflecting the changes necessitated by the problems discussed above, the revised version omits some of the superfluous details of the original passage, and replaces two passive constructions by their active equivalents:

After the examination, the school sends the scripts to the examiner. He marks them and writes the marks on a form which he sends to the examining board. At the examining board the marks are typed into the computer system. This is a slow and tedious process.

.....

"As you write your subalgorithms below, define any local variables used. Do not use global variables. State if the subalgorithm is a function or a procedure.

(a) Write a subalgorithm for the submodule .." (22)

This short extract exemplifies one of the most important issues facing the examiner; namely, where to place amplifying instructions. There is a constant possibility that the overanxious or poorly-prepared and inexperienced candidate may start writing as soon as he has identified the main task and thus fail to pay sufficient or any attention to the important qualifying information. Mindful of this fact, the examiner may decide to place this information at the start of the question. The problem with this, however, is that such information functions as a forward (cataphoric) reference and has to be held in abeyance by the candidate until the main instruction is encountered. Moreover, if the reference is inexplicit, there is a danger that on the first reading no connection will be made to the information which follows.

For two inter-related reasons, there is a distinct possibility that candidates will not correctly identify the cataphoric reference of the word 'below' in the above extract. Firstly, because the unnecessary use of the word 'your' may imply to some candidates that the algorithms already exist or have already been referred to in the half-page introduction to the question. Secondly, because the word 'below' may be initially be read as qualifying 'write' rather than qualifying 'subalgorithm'. In other words, candidates might at first assume that they have to write the subalgorithms in the space at

the foot of the question page. So assuming the pre-position of the amplifying information is to be retained, the cataphoric reference needs to be made more explicit.

One further improvement that can be made is to replace the polysemous conjunction 'as', which could be interpreted as meaning that the definition of the local variables is an ongoing or repeated process, whereas this needs to be done once only.

When you answer questions (a), (b) and (c) below, define any local variables you use. Do not use global variables. State if the subalgorithm is a function or a procedure.

(a) Write a subalgorithm for the submodule.

.....

"The ten pairs of integers below indicate points earned for the home team and the opponents, respectively.

H	o
6	14
10	9

Clearly, the need to provide students with sufficient background information to understand a task often conflicts with the desire to avoid over-long questions. The decision as to which should take preference, detail or brevity, must be made case by case. However, it seems that the wish to be brief has been taken too far in the above extract, which is the opening sentence of the question. In particular, the candidate may be puzzled by the use of 'the' to qualify 'home team' and 'opponents'. This type of cataphoric usage of the definite article is not uncommon among such diverse writers as poets and journalists who wish to create a dramatic effect and capture the reader's immediate attention - cf. "*The boy stood on the burning deck* .." and "*The bullet shattered the window and lodged in the wall above the woman's head.*" The tacit agreement underlying such a convention is that the questions "*Which boy?", "Whose bullet?", "Which woman?",* etc. will be resolved in the text that follows. The problem with the examination extract above, however, is that the question "Which teams?" is not further clarified. Indeed, there is no reference at all to the type of sport or game that is being played.

As well as removing this potential source of confusion, the extract can be improved in two further ways. Firstly, if the phrase 'earned for' is replaced by the more natural 'won by'; and secondly, if the barely legible, hand-written headings to the two columns are replaced by proper titles, thereby removing the need for the word 'respectively';

8. The pairs of integers below show the points won by the home team and their opponents in a series of ten games of baseball:

HOME TEAM	OPPONENTS
6	14
10	9

.....

A particular kind of reference is effected by the use of anaphoric nouns, such as *'approach'*, *'purpose'*, *'problem'*, which refer back in a general way to ideas or abstractions which have often been explained over large chunks of text (Francis, 1987). As such, they are different from words like *'algorithm'*, *'program' 'flowchart'* etc., which refer unambiguously to concrete entities. It is precisely because of their generality and potential for ambiguity that such nouns need to be used with care in examination questions.

"Produce an algorithm for a sub-program which would process and calculate the charge for a single call. Your solution should ideally include all four built-in functions described below."

The use of the word 'solution' in the second sentence may make this short extract stylistically more agreeable, but candidates are probably better served by a direct repetition of the word 'algorithm'. Furthermore, the word 'ideally' may be difficult for some candidates. The instruction is probably more intelligible as follows:

Write an algorithm for a sub-program which would process and calculate the charge for a single call. If possible the algorithm should include all four built-in functions described below.

.....

There is no suggestion that the short original extract above would cause great difficulty. Indeed, the more serious problems tend to occur when the referent is a large chunk of text situated at some remove from its anaphoric noun. Such problems are difficult to analyse briefly, however, since they require the reproduction of whole questions or long sections of them. The following two short extracts must suffice as exemplification:

"Instead of doing this problem for strings of length 3 taken from a 3-letter alphabet, one could just as well use strings of length m from an alphabet consisting of n letters. On the other hand, instead of using a recursive procedure, we (sic) could solve this problem straight forwardly (sic) using an m-fold structure.

Explain why the solution with a recursive approach would be advantageous in the above situation." (20)

Comprehension of the task to be performed is made difficult because of the use of three anaphoric nouns in the last sentence - 'solution', 'approach' and 'situation' - none of which have explicit referents. The instruction is much clearer as follows:

Why would it be better to use a recursive procedure to solve the above problem?

.....

"Both the line printer and the VDU (Visual Display Unit)/CRT (Cathode Ray Tube) are important devices for providing information. Such information may by (sic) provided during programming development through a computer, or as output from the program.

Give two suitable applications of each type of device (line printer and VDU/CRT) describing in each case why it is appropriate for the purposes given." (4)

The candidate would be forgiven for not realising that the anaphoric noun 'purposes' in the phrase 'for the purposes given' is an exophoric (out of text) reference, not an anaphoric one. In other words, the purposes have not already been given in the preceding text, but are to be given by the candidate in his answers. The instruction should be revised to avoid this ambiguity and also to rectify the problem caused by the missing comma before 'describing' :

Give two suitable applications of each type of device. In each case state why the device is appropriate for the application you have chosen.

5.5 Organization

The organization or structure of an examination text plays a crucial role in helping candidates to understand exactly what they are required to do. The responsibility of the examiner, then, unless he or she intends part of the task to involve the correct identification of what is to be done from a mass of extraneous or ambiguous information, is to formulate instructions in a clear, concise and consistent manner. One way this can be done is by itemization, whereby multiple tasks, embedded in a lengthy sentence or paragraph, are separated out and listed one by one. The following extract, for example, is probably more immediately intelligible in the itemized form shown beneath it.

"(b) Write a correct algorithm for the original problem cited in (a) and indicate using output statements the result of the search: if the list is empty, if the value of searchitem is in the list or if the value of searchitem is not in the list." (21)

Revised Version:

(b) Write an algorithm to search a linked list, as explained in (a).

(c) Use output statements to show the result of the search,

(i) if the list is empty

(ii) if the value of *searchitem* is in the list

(iii) if the value of *searchitem* is not in the list.

As well as the itemization of tasks, the original extract has been improved by attending to the other minor problems it contains. So, the embedded phrase `'using output statements'` has been relocated, the infrequent word `'cited'` replaced by the more common 'explained', and the thin punctuation filled out.

.....

Another way that candidates can quickly and accurately identify the tasks they have to perform is if there is some consistency in the way that the various sections of the questions are ordered. A question in the IB Computing Studies paper, for example, may consist of the following different parts:

1. a type of 'scene-setting' or preamble in which the necessary background information describing the computer application is set out;

2. a number of technical assumptions that the candidate must take into account in performing the tasks;

3. the tasks themselves;

4. information amplifying the tasks - for example, advising the candidate on the form of his answer.

It is probably very helpful if the candidate can identify these different sections quickly and accurately. He is more likely to be able to do so if they follow a consistent order across examination papers and/or are clearly labelled. The whole question which is analysed in part 5.11.2 of this paper gives an example of how the absence of transparency in a question's organization can add to the time the candidate needs to understand what he has to do.

5.6 Punctuation

Comprehensibility is aided by precision of language, and one way that precision is achieved is by accurate and appropriate punctuation. There seems to be a trend in British writing towards reducing punctuation to a minimum, especially in popular journalism and fiction where the aim is a 'snappy', fast-moving style, with the writer relying on the reader to, as it were, supply the punctuation himself.

It is a different matter in examination texts; here it is essential for the writer to make it as easy and quick as possible for the reader to reach a clear and complete understanding of the task he has to perform. This can be achieved in particular by helping him to identify the clauses of complex sentences by separating them off with commas. If this is done, the reader is likely to have fewer problems in efficient intrasentential chunking, (Nuttall, 1983). An example of how incorrect chunking can be caused by under-punctuation is shown in the extract below, which is one fifty-word sentence.

"The system to be designed must allow for future development and an estimate has been made that a twenty per cent expansion in the number of items and the number of manufacturers with a fifty per cent growth in the number of retail outlets should be possible in the system." (1)

The length of the sentence alone places a great strain on the reader's short-term memory, and the absence of any punctuation compounds the problem. The missing comma after 'development' could lead the candidate to read 'an estimate' as the second object of the verb 'allow for', i.e. that the string of words 'allow for future development and an estimate ..' is one chunk. Of course, this hypothesis will be rejected as soon as the candidate reads on, but this may mean a regression in order to reinterpret the meaning of the first part of the sentence. Similarly later in the sentence the string 'the number of manufacturers with a fifty per cent growth' may be erroneously decoded as one chunk.

In this case, the comprehensibility of the passage can be greatly increased if the information is itemized:

The system to be designed must be capable of the following expansion:

(i) a 30% increase in the number of items,

(ii) a 20% increase in the number of manufacturers,

(iii) a 50% increase in the number of retail shops.

.....

There is a very clear tendency in the IB Computing Studies Papers to 'under-punctuate', either by omitting commas as in the example above or by running clauses together instead of starting a new sentence. As the following extracts show, this feature often produces texts which are more difficult and take longer to interpret than necessary.

"If you assume that the names of countries are stored in 4-character blocks what effect does this have on preparing a suitable program to carry out the sort completely." (1)

The missing comma after 'blocks' and the missing question mark at the end of the sentence make this question more difficult to process than it need be. The task is probably more easily comprehensible expressed in two sentences, with the phrasal verb 'carry out' replaced by its simple synonym 'do':

Assume that the names of the countries are stored in 4-character blocks. What effect does this have on preparing a suitable program to carry out the sort completely?

.....

"The time will be given in whole minutes, beginning with 0 at midnight, then running to 1439 thereafter 0 follows again since a day has 1440 minutes."

The missing comma after 1439 and the formal word 'thereafter' could make this passage a little difficult for some candidates. Two separate sentences may make it easier for them to process the information more quickly:

The time is given in whole minutes, beginning with 0 at midnight and running to 1439. Since a day has 1440 minutes, 1439 is followed by 0 again.

.....

The use of a present participial clause (PPC) to amplify an instruction is another very common feature of the IB Computing Studies papers. Sometimes the PPC is separated off by a comma, which tends to make the text more comprehensible. However, very often no comma is used, which can result in the erroneous interpretation that the amplifying instruction is in fact a reduced, post-modifying relative clause. The following extracts are just a few of the many examples of this problem; they are followed in each case by a suggested improvement. The improvement is effected either by inserting the conjunction 'and' or by starting a new sentence, and by replacing the present participle with an imperative or an interrogative:

"Suggest another method, other than magnetic strips, which might be used for input of the required data (competitor number/finishing time) specifying the hardware that would be needed." (5)

What method other than magnetic strips could be used for input of the required data (competitor number/finishing time)? What hardware would be needed?

.....

"Describe a 'real time' application explaining why it is 'real time'." (12)

Describe a 'real time' application. Explain why it is 'real time'.

.....

Explain briefly the difference between analog and digital measurements giving an example of each." (12)

Explain briefly the difference between analog and digital measurements, and give an example of each.

.....

Given q <-- false, A <-- 7, C <--1, evaluate the expression by specifying the result of this expression and by locating the appropriate line in the truth table in part (i) using clear labelling." (21)

Given q <-- false, A <-- 7, and C <--I, evaluate the expression by stating its result and by indicating the appropriate line in the truth table in part (i). Use clear labelling.

.....

The following extract has a different example of how a lack of punctuation or special marking can cause problems:

"Assume all variables are of type real except the variable depends (number of dependants), which is integer, and the variable validdata, which is Boolean." (22)

In this case the variable name ('depends') could at first be read as a verb. An erroneous decoding of this kind would not happen if the variable were marked in some way. Indeed, it would be highly desirable if some convention could be adopted for the consistent treatment across examination papers of the names of variables. In some papers these are given in bold type, others have an initial capital letter, and still others have all letters capitalized. The same applies to numerical variables which are identified variously as ith, ith, 'i'-th, ath and nth. For example:

"If a move takes you outside to the right of the square in the ith row, place the integer in the ith row at the left side." – (18)

.....

Punctuation is not subject to the same strict rules that apply to other areas of written language such as spelling or syntax. Punctuation, particularly in the use of the comma, is more a matter of personal choice and style. As Quirk (1985) says;

> *There is a great deal of flexibility possible in the use of the comma, in its presence or absence, or in its replacement by other marks;*

Often the most that can be said of unconventional punctuation, as in many of the examples above, is that it is makes the reader's task harder than it need be. Nevertheless, there are certain rules about how to punctuate which every writer should follow. One such rule is that non-defining relative clauses and other 'included units' (Quirk, 1985) must be separated off by commas. Unfortunately this is a rule that is frequently broken in the IB Computing Studies papers. For example:

"A particular program, written in a high level language fails a number of times during its development." (3)

and:

"The purpose of this subalgorithm is to compute the gross weekly pay for an employee which includes regular pay and overtime pay." (22)

In the first example above, a comma is needed after 'fail'. In the second, not only is the non-defining relative clause starting 'which ..' not marked off by a comma, but it is also separated from its antecedent 'pay'. This makes processing the necessary information harder than it might be. The passage is probably clearer if the relative clause is expanded into a sentence and put into brackets at the end:

The purpose of this subalgorithm is to calculate the employee's gross weekly pay. (Gross weekly pay includes regular pay and overtime pay.)

.....

"Design in flowchart form or using pseudocode, the program which would run at the end of each day to process payments made on that day to carry out tasks (iii), (iv), (v) and (vi)." (8)

As it stands the sentence above is wrongly punctuated - the embedded nominal group ' in flowchart form or using pseudocode' must either be marked off by two commas or by none; in fact, it is better relocated at the end. Furthermore, the tacking together of the two infinitive clauses 'to process payments ..' and 'to carry out tasks ..' without punctuation or a conjunction makes the passage very difficult to understand. It is probably easier as:

Design a program which would run at the end of each day. The program should process any payments that were made that day, and it should carry

out the tasks in (in), (iv), (v) and (vi) above.

Write the program in pseudocode or flowchart form.

.....

"Design in flowchart form or using pseudocode a subprogram which will store in an array C, the sum of the integers held in two arrays, A and B." (12)

This time the embedded unit has not been marked off by commas. This is acceptable, although, as we have seen sentences containing embedded strings are often more difficult to decode. The problem, however, lies with the comma after 'array C', which has separated the verb from its object. This is incorrect punctuation, as Quirk (1985) makes clear:

> *With regard to the central clause elements S, V, O and C, there is a strict rule that they cannot be interrupted by punctuation.*

The passage can be re-ordered and slightly reworded in order to remove the faulty comma:

Design a subprogram which will store the sum of the integers in arrays A and B in a new array, C. Write the subprogram in pseudocode or flowchart form.

.....

There is a similar problem of interruption of subject and verb in the next extract:

"An * at the end of an assembly code instruction or the beginning of a line, indicates that all the following characters on that line are comments." (26)

It is possibly the distance of the head 'word' of the subject, ('*'), from its predicate ('indicates ..') that has occasioned the faulty use of the comma after 'line'. To make the information in the sentence more immediately transparent, the word 'at' can be repeated in order to signal clearly to the candidate that the nominal group is not yet at an end:

An * at the end of an assembly code instruction or at the beginning of a line indicates that all the following characters on that line are comments.

.....

Another clear-cut rule of punctuation is that only sentences should be terminated by a full stop. This

rule is broken in the following extract:

"Using the tree diagram and the array structure, TREE, composed of fields: Location, Data, LeftChild, and RightChild; and fields Root and Avail. Location is the relative address of the left and right children of the node .." (16)

The full stop after Avail will compel the candidate to reject the probable initial hypothesis that the first words of the sentence constitute a dependent clause, and will cause him to regress in order to make sense of what he has just read. 'Using' should be replaced by 'Use' to avoid this unnecessary increase in processing time.

.....

The many examples above show that punctuation has a great deal of influence on the comprehensibility of texts - as Kahn (1985) says, ambiguity thrives on careless punctuation. In examination texts in particular it is vital that punctuation is not careless but precise and accurate.

5.7 Naturalness

The choice of an unexpected word or an unusual syntactical form is a stylistic device often employed for effect by writers who wish to slow the reader down and make him conscious of the language being used. The use of such devices in an examination, however, would seem inappropriate, since the focus should always be on the message (i.e. on the tasks to performed) and not on how the message is conveyed (i.e. the language in which the tasks are expressed). Any unnaturalness of language may prove an undesirable distraction to the candidate. But what is naturalness in language?

McCarthy (1988) defines it as:

.. a quality of utterances that renders them liable to acceptance or rejection as 'natural' over and above decisions as to their grammatical, lexical or semantic well-formedness.

In other words, a sentence can be perfectly correct grammatically and lexically, but still appear rather odd. This begs the question as to the nature of the sentence's oddity, and whether it will seem equally odd to the professional linguist as to the man on the street. Indeed, as McCarthy admits, we do not, as yet, have convincing 'models' of naturalness in language. However, he does suggest some criteria by which we can make reasonably reliable, objective judgments about any given utterance. The most relevant criterion from the point of view of this study is called 'preferred sequencing', which encompasses word order, collocation and colligation.

Although not all of the extracts which follow violate the criterion of preferred sequencing, each of them appears odd or inappropriate in its own right and by virtue of its unnaturalness may distract the candidate. This is sufficient reason to suggest it should be amended, and when the unnaturalness results in ambiguity there is even more justification.

"It is now required to store a third stack in the array A.

Briefly describe how you could achieve this. You are not required to write programming language statements for each operation." (11)

The impersonal passive expression "it is required to" + verb, meaning something has to be done, is very unusual. It possibly results from a faulty analogy with a phrase containing an adjective such as 'it is necessary to ..'. A more concise and elegant way of expressing the instruction is:

Describe how you could store a third stack in the array A. (You do not need to write programming language statements for each operation.)

.....

"A computer has a hard disk on which are stored files belonging to many users." (12)

The unnatural word order in the relative clause reverses the subject and verb. The sentence is more natural as;

A computer has a hard disk which stores files belonging to many users.

.....

"Not all the RAM is available to the user. Why is this? Suggest for what it is being used." (15)

The attempt to avoid finishing the sentence with a preposition has lead to an unnatural, possibly ambiguous instruction. Better is:

What do you think it is being used for?

.....

"To what value do the rows, columns and diagonals sum?" (18)

To 'sum to' is an unusual collocation, and the word 'value' is superfluous. (The issue of whether one sum or three sums are required will not be pursued.) The question can be more naturally expressed as:

What is the sum of the rows, columns and diagonals?

.....

"A file of 1200 records each 200 characters long is stored on the tape with a blocking factor of 5.

Find

(i) the space occupied by the file

(ii) the time taken to load the file assuming single buffering." (18)

The instruction 'find' is rather odd and may initially confuse some candidates as it seems to contain a misleading element of random searching. More natural is a direct interrogative. The new version, with corrected punctuation, reads;

A file of 1200 records, each 200 characters long, is stored on the tape, with a blocking factor of 5.

(i) How much space is occupied by the file?
(ii) Assuming single buffering, how long does it take to load the file?

.....

" 3. You are to consider all possible 3-letter combinations (strings of length 3) consisting of the letters 'A', 'B', 'C': e.g. AAA, BCA, etc. The same letter may occur more than once.

(a) How many different combinations are there using strings of length 3 and 3 letters?

(b) List by hand all the strings of length 3 in alphabetical order." (21)

'You are to' is an unnatural way of setting an examination task - this is usually done using the imperative form. Furthermore, it is unclear what is intended by the instruction in part (b), 'list by hand'. To state that something should be listed by hand is unusual and implies that it should not be listed by other method likely to be chosen by the candidate. If this is indeed the case here, then that other method should be explicitly forbidden. This would avoid giving the candidate cause to stop and ponder exactly what is required.

Another slight problem is the separation of 'list' from 'in alphabetic order', which could lead to 'all the strings of length 3 in alphabetic order' being read as a single nominal group. To avoid these difficulties the instruction could be radically rewritten as follows:

3. (a) How many different strings of length 3 can be made from combinations of the letters A, B, and C?

Note: i. Each string must contain 3 letters. ii. A letter can be used more than once in a string; e.g. AAA, BBC.

(b) Write an alphabetical list of all the 3 letter strings.

.....

The article system in English is highly complex and allows a certain degree of individual choice. There are, however, a number of widely-accepted rules of use and when these rules are contravened the results are unnaturalness or even ambiguity of language. Although most the following examples of unnatural article use are unlikely to cause serious confusion, they may distract the candidate to such an extent that he feels the need to regress and reprocess what he has just read. This is undesirable and is sufficient justification for the suggested amendments.

"A large city is a very popular centre for tourism. Each year, it attracts many visitors. The Local Tourist Board, who are (sic) responsible for ensuring the success of local attractions, feel that a number of computer operated 'guides' would help visitors wishing to visit the attractions." (6)

The problem here is the meaning switch between sentence 1 and sentence 2. The reference to 'city' in the first sentence, which is also the first sentence of the question, will probably be understood in its generic sense, i.e. as referring to any or all large cities. There is initial confusion, then, when the candidate reads on and realizes that reference is being made to a specific city with a specific problem, to be solved by computerization. This confusion can be avoided if the city, real or fictitious is named, or if the preamble makes the switch from general to specific more explicit as follows:

Cities are very popular tourist centres and attract many visitors each year. The Tourist Office of a large European city, which is responsible for the success of local attractions, feels that ..

In fact, much of the information in the first part of the question is superfluous and the task can be understood perfectly well if it is omitted:

Assume that the tourist office of a large city plans to set up a number of computer-operated 'guides' to help tourists who want information about the city's attractions.

.....

In the next question to be considered candidates have to organize data about fifteen summer camps. This is part of the question:

"Using the definitions above and defining any additional variables, write subalgorithms that would do the following tasks.

(i) Determine a camp which spent the most money on computers.

(v) Determine an activity which was the most expensive of the three activities across all camps.

(vi) Determine an activity which was the most popular in terms of camper participation." (24)

(Each of the remaining tasks (vii) - (iv) starts with the instruction 'determine'.)

The indefinite articles imply that there could have been more than one camp which spent the most money on computers, or more than one activity which was the most expensive etc., but this is clearly absurd. The instructions can be expressed more precisely as follows:

Write subalgorithms that would determine (i) - (vi) below. (Use the definitions above and define any additional variables.)

(i) The camp that spent the most money on computers.

(ii) The most expensive activity.

(vi) The most popular activity.

.....

It has been noted that nonnative readers' comprehension of texts may be improved by the inclusion of redundant features such the 'wh..' + be construction in relative clauses and the filling out of other types of ellipsis. However, there is another type of redundancy which is less helpful, namely the presence in a text of unnecessary material. Since in many questions in the IB Computing Studies examination the candidate has a great deal to read, sometimes as much as two full A4 printed pages, it seems reasonable to suggest that there should be no superfluous information. This is especially so in cases where this information may only serve to create ambiguity. Of course, the suggestion does not allow for the possibility that the examination board believes that part of the process of separating the good candidate from the poor one consists in seeing who can extract the necessary information from long and difficult texts under considerable time pressure.

In the following extracts the superfluous parts of the text are shown by underlining.

"Discuss briefly other changes that have happened or are likely to happen in the modern office through the introduction of microtechnology, based on information you have gained during your course." (2)

The problem here is that the careful candidate would be justified in asking himself whether any knowledge about office microtechnology that he has from gained from sources other than the classroom is admissible in his answer. To avoid any possible ambiguity, the underlined words should be omitted.

.....

There is a similar problem in the next extract:

"Write a program in any high level language with which you are familiar to assist a currency exchange bureau." (2)

It can reasonably be assumed that candidates who want to achieve good marks will write the program in a language with which they are familiar rather than in one with which they are not familiar, so the clause is superfluous.

.....

"A number of different types of "error" may result from the arithmetical processing of a computer program. From your knowledge, give two examples of errors that may arise when the program is run." (4)

The phrase is unnecessary since the candidate cannot be expected to give examples which are not from his knowledge.

.....

A similar comment can be made about the following extract:

"Explain, from your experience, how the computer might react to each type of error when you were entering the program on the keyboard." (4)

In this case, the pedantic student who has only indirect knowledge, but no direct personal experience, of entering faulty programs on a keyboard might be confused as to whether his answer is permissible.

.....

"The system must be able to carry out, amongst others, the following operations:

(i) record the issue of a book subject to the rule that no one may have on loan more than four books at a time;" (6)

The embedded phrase adds nothing of importance to the instruction. It is also potentially problematic because of the rather opaque cataphoric reference of 'others' and the uncommon word

'amongst'. In fact, the extract shows a considerable increase in clarity if a large portion of the first sentence, including the offending phrase, is omitted:

The system must be able to:

(i) record the issue of a book, subject to the rule that no-one may have more than four books on loan at a time,

.....

"The college has an advisor/counsellor to the students who helps the students with their choice of subjects and progress." (15)

There seems to be no good reason for the alternative to 'advisor' or for the repetition of 'students'. After excision of these words, the passage reads more simply as:

The college has an advisor who helps the students with their choice of subjects and with their progress.

.....

"Give at least two advantages of having a hard disk as opposed to a floppy disk." (15)

"Give at least one difference and one similarity between the two data structures in each part." (16)

Presumably, candidates have nothing to gain by writing more than the minimum required, in which case the phrase 'at least' is superfluous in the above questions and should be omitted to avoid candidates wasting precious time.

.....

"Consider the broader problem of temperature control. Assume that we wish the computer not only to monitor the temperature but also to control the temperature in the office. Keeping the office warm is not a problem due to very effective solar heating units and backup heating units. However, keeping the office cool is the problem. The office has an air-conditioning system which can be used to keep the temperature in the office around a given temperature, T. Each day at 0700 hours, the computer should start up the system by calling an existing subalgorithm START-UP. At 1700 hours, the system should be shut down by a call of the existing subalgorithm SHUT-DOWN. " (20)

The above extract is about one quarter of a long, closely packed question. Particularly in this case, therefore, it is desirable that the candidate's reading load is kept to a minimum. This can be done by a fairly ruthless excision of all superfluity; including, for example, the reference to the two types of heating unit in the sentence starting 'Keeping the office warm . . '. This information is superfluous to an understanding of the task to be performed:

Assume that we wish the computer to control the office air-conditioning system and that we want to keep the office around a given ideal temperature, T. Each day at 0700 hours the computer should start up the air-conditioning' system by calling an existing subalgorithm START-UP. Each day at 1700 hours it should shut down the system by calling an existing subalgorithm SHUT-DOWN.

.....

"Explain in one sentence what this set of instructions accomplishes.' (22)

The problem here is that there is no upper limit to the length of a sentence. If a brief answer is essential, then a maximum number of words should be stated; if not, the phrase should be omitted. Furthermore, 'accomplishes' is unnecessarily difficult and could be replaced by 'does'.

5.9 Ambiguity

Difficulty may be caused not only when words are not known or when complex syntax cannot be untangled, but also in cases where the text is open to more than one interpretation, i.e. where it is ambiguous. We have already seen how ambiguity can result from faulty punctuation or unclear reference, but it can also arise in other ways. Even the seemingly innocuous words 'and' and 'or' can cause confusion and need to be used with great care, particularly in instructions. The following extracts are potentially difficult because of the ambiguity of the conjunctions they contain.

For example, it is not clear if the following question requires one answer or two:

"Give a further example of the use of RAM and ROM in a microcomputer." (4)

Assuming two separate answers are needed, the instruction can be made less ambiguous by itemization:

Give a further example of the use in a microcomputer of 1.) RAM and 2.) ROM.

.....

It is unlikely that any real difficulty was caused by the above example, but it is a different matter in the next two extracts:

"Discuss the advantages or disadvantages of using an interpreter or a compiler in (i) the development of the program (ii) the subsequent use of the program." (6)

There is a multitude of possible interpretations to this instruction. For example, for each of (i) and (ii) does the candidate have to state the advantages AND disadvantages or the advantages OR disadvantages? And does he have to do this both in respect of the interpreter AND the compiler or for one of them only?

"Describe two methods that would help make this system withstand improper or incorrect use." (15)

Similarly, it is not clear whether this question requires the candidate to give two answers about improper use OR two answers about incorrect use, or one answer about improper use AND one answer about incorrect use. On the assumption that the first of these is the correct interpretation, the question could be reformulated:

EITHER:

Describe two methods that would help make this system withstand improper use.

OR:

Describe two methods that would help make this system withstand incorrect use.

.....

"You may find the list of variable names below and their descriptions helpful in formulating your answer. You may wish to use additional variable names and/or you may find you do not need all the names supplied." (14)

The 'and/or' conjunction is often problematic and needs to be used carefully. If the construction in this passage is 'unpacked', the following sentences result:

a.) You may wish to use additional variable names and you may find you do not need all the names supplied.

b.) You may wish to use additional variable names or you may find you do not need all the names supplied.

Sentence b.) is acceptable, but clearly sentence a.) makes no sense; if the candidate does not need to use all the names supplied, he is unlikely to need to use additional ones. The instruction is clearer as follows:

You may find the following list of variable names and their descriptions helpful in writing your answer. If you need to use extra variables, you should make up your own names for them.

.....

"Some of the processing is carried out interactively and some in batch mode. State which of the activities 1. to 4. above is more appropriately carried out in either interactive or batch mode, giving reasons for each answer." (19)

The either/or construction is incorrect in this extract because it implies that at least one of the activities is not appropriately carried out in either batch or interactive mode. (If you are asked to identify which of ten objects on a table is either a fruit or a vegetable, there is a clear implication that there are other types of objects on the table. You point to the banana, potato, lemon etc. but do not point to the spoon or the paintbrush.) The reformulated instruction below is less ambiguous:

Some of the processing is carried out interactively and some in batch mode. For each of the activities in 1. to 4. above, state if it is more appropriately carried out in interactive mode or in batch mode. Give a reason for each answer.

5.10 Instruction Words

Some of the types of problem that have been discussed so far, such as embedding or under-punctuation, may be considered to be less serious, particularly where these occur in a preamble to the task(s) or in supplementary information. In other words, although such difficulties may cause an interruption of the reading flow or create a temporary barrier to comprehension, the candidate will probably, after some effort, arrive at a correct or approximate understanding of what he has read, and even if he does not, his ability to answer the question set may not be prejudiced.

It is a different matter, however, in the case of the direct instructions to the candidate about the task he has to perform, where any inaccurate or less than complete understanding of the question is likely to have a direct bearing on the quality of answer. It is crucial, therefore, a.) that instructions to perform a task are clearly seen as such, and b.) that the instructions themselves are expressed in absolute1y unambiguous 1anguage. This is true for any examination, but particularly for one which is taken by large numbers of nonnative speakers. Unfortunately, the following analysis will show that many of the instructions in the IB Computing Studies papers do not meet these minimal requirements. The extracts contain a variety of the problems which have already been discussed, but common to all of them is the

potential difficulty or ambiguity which arises from the instruction word itself.

In a large number of cases the problematic instruction words, which are usually in the imperative form, can be omitted altogether or replaced by an interrogative.

"Give a brief outline of the data files that you think will be needed indicating the probable size of each." (1)

The instruction "Give an outline . . " is not particularly transparent, and could be interpreted as requiring some sort of illustration. In addition, the passage suffers from the familiar problem of failing to mark off the clause containing amplifying information. In fact, the task is probably more intelligibly expressed as two interrogatives:

What data files do you think will be needed? How long will they (probably) need to be?

.....

"Outline briefly the essential hardware which you consider would be needed for the carrying out of this task." (5)

In this extract, too, the instruction 'outline' is unnecessary and potentially ambiguous. The question can be further shortened by removing the tautologous combination of 'essential' and 'needed' and by reformulating the unnatural nominalization 'the carrying out of this task' as a verb phrase. The revised version reads:

What hardware do you think would be needed to perform this task?

(Phrasal verbs such as "carry out" are notoriously difficult for nonnative English speakers. For this reason it has been replaced by the word "perform" in the suggested improvement above.)

.....

"In a single example, give a pictorial or programming language representation using all four parameter types." (13)

The problem here results from the attempt to keep the question short. Firstly, the candidate is not told explicitly what the example should consist of. Secondly, the word 'representation' does not collocate with either of the two alternatives whose function it is to combine. In other words, it is equally unnatural to instruct the candidate to give a pictorial representation as it is to tell him to give a programming language representation. The text needs to be expanded to make the candidate's task more explicit:

Give an example of a subprogram which uses all four parameter types. (You

may draw a diagram as your answer, OR you may write your answer in a programming language of your choice.)

.....

"Describe the output that should be generated if the input is:

(i) −507 −33 0 192 42 −999 .." (18)

The instruction 'describe' may mislead the candidate into thinking he has to give a more elaborate answer than is actually necessary. In fact all he has to do is state what output will be produced. A simple interrogative is much less ambiguous:

What output should be generated if the input is.

(i) -507 -33 0 192 42 -999

.....

(d) "Which solution is more efficient in terms of storage of binary digits? Which solution is more efficient in terms of the programmer-generated instructions executed? Support your answers quantitatively."

The instruction in the last sentence is not particularly transparent. It could be expressed more naturally if the question 'and by how much?' were added to the preceding two sentences. Furthermore, the extract has another example of an anaphoric noun, 'solution'. Instead of using this word, which has not occurred in the preceding text, it would be clearer to refer to exactly what was required in the previous two questions, namely a subalgorithm.

(d) (i) Which subalgorithm is more efficient in terms of the storage of binary digits, and by how much?

(ii) Which subalgorithm is more efficient in terms of the programmer-generated instructions executed, and by how much?

.....

Swales (1982) points out that the meanings of certain verbs like *discuss*, *describe* or *compare* take on conventionalized meanings in examination papers. The instruction *'illustrate'*, a further example of this so-called 'examinese', is used very frequently in the IB Computing Studies papers, where, in most cases, it has the conventionalized sense of 'give examples in support of your answer'. A potential difficulty for candidates, however, is that the word is also used in its more everyday sense of ' draw a diagram', and in yet other cases it seems to mean something else entirely. Consequently, the candidate has to spend time puzzling out exactly what the instruction requires him to do. The following examples illustrate (!) this lack of consistency.

"Explain the difference between logical programming errors and syntax errors. Illustrate your answer by giving examples of two errors of each type .." (5)

'Illustrate' is being used in the conventionalized sense here, but since it is in any case followed by reference to the need for examples, it is unnecessary and better omitted altogether:

Explain the difference between logical programming errors and syntax errors. In your answer you should give two examples of each type of error.

.....

"Write a brief note illustrating the use and importance of each of the following in a computer system:

a) an operating system

b) high level language compiler (sic)

c) an editor." (5)

The first part of this instruction is superfluous and in any case a candidate may be unsure exactly what is meant by 'a brief note'. The question can be expressed interrogatively:

What is the use and importance of each of the following in a computer system?

a) an operating system

b) a high level language compiler

c) an editor

.....

In the following extract 'illustrate' is used twice:

"(a) (i) What does it mean for data to be consistent?
(ii) Using the employee record above illustrate the lack of data consistency.

(c) Using the recent record changes illustrate where data redundancy occurs and also the lack of data integrity." (19)

The problem in part (a) is that a diagram showing an employee's record is shown on the examination sheet, and the instruction `'illustrate'` may mislead some candidates into thinking they have to reproduce this diagram and draw arrows etc. to show the inconsistency in data representation. In fact, `'illustrate'` here simply means 'point out'. To avoid any confusion instruction (a) (ii) could be rewritten as an interrogative:

What is inconsistent about the data in the employee record in Figure 1 above?

The similar ambiguity of the word `'illustrate'` in task (c) is compounded by the odd use of `'recent'` to mean 'as shown in the second diagram above' and by the awkward conflation of two instructions using the word 'also'. A suggested improved version is:

Look at the changed employee record in Figure 2 above and state:

(i) where there is data redundancy,
(ii) where there is a lack of data integrity.

On a general note, where there is more than one chart or diagram in an examination question, as in the above extract, it is helpful to the candidate if these are labelled and referred to explicitly by name and number; e.g. Table 1, Figure 2, etc.

.....

`Illustrate what output the program would produce if left uncorrected and the numbers 40, 25 were entered as data."` (3)

In this question the instruction `'illustrate'` serves no useful function whatever, and again may mislead the candidate into thinking that a special type of answer is required. A simple interrogative is much less ambiguous. In the revised version below, the ellipsis has been 'filled in' and the slightly simpler first conditional has been used:

What output will the program produce if it is left uncorrected and the numbers 40, 25 are entered as data?

.....

`"Given the data below for CUSTARRAY, graphically illustrate the contents of ORDEREDID."`

In this case, a table or chart is what is needed. Possibly because the instruction `'illustrate'` has been used so often to mean 'give examples', it has been felt necessary to qualify it in this case by the word 'graphically'. Unfortunately, this may have the effect of making the candidate more rather than less confused about exactly what he has to do. '.. fill in a table with the contents of

ORDEREDID ' is better.

.....

Another common instruction word in these examination papers is 'indicate', which takes on the conventionalized meaning of 'state', as in the following examples:

"Indicate which non-computer staff should be involved at the different stages of the development of the payroll application." (5)

"Indicate the hardware that you think might be most suitable for this application giving reasons for the choice." (6)

In most cases, however, 'indicate' is unnecessary because the instruction can usually be more simply expressed interrogatively. So, for example, the first extract above could read:

Which non-computer staff should be involved at the different stages of the development of the payroll application?

The second extract, as well as being expressed interrogatively, would also benefit from revision in order to correct the faulty punctuation and to remove the odd use of the definite article to qualify 'choice'. Moreover, it may also be helpful to the candidate to tell him how many reasons he should give:

What hardware do you think would be the most suitable for this application? Give (two) reasons for your choice.

In the following extract 'indicate' is used in its more everyday sense of 'show' or 'point out':

"Label all components and indicate clearly on the diagram the following .." (15)

So the difficulty with ' indicate' is not necessarily in any one question, but results from a lack of consistency across different questions, which requires the candidate to work out each time exactly what is needed.

.....

The term 'give' is another example of 'examinese'. In most cases there is no objection to this instruction, which will be interpreted by candidates as meaning 'state' or 'write', as in "Give

three advantages of linking the computers into a network." (22), or "Give a general definition of floating point representation." (21). Nevertheless, instructions which are expressed in more natural occurring language are to be preferred, since then there is less possibility of ambiguity arising. For example, "Give two types of errors (sic) that can occur in binary arithmetic" (19) is less natural than "What are two types of error that can occur in binary arithmetic?"

'Identify' is another word which is used in the IB Computing Studies examination questions to carry a meaning different from its everyday one. Like 'give' it tends to mean simply 'state' or 'name', and it too can often be replaced by an interrogative. So, for example:

"For each pseudocode fragment, identify what is wrong with the logic." (21)

is more natural as:

What is wrong with the logic in each pseudocode fragment?

.....

Apart from the instruction 'identify', the following passage is difficult for two main reasons: the embedded phrase 'other than data independence' in the long first sentence, and the ellipsis and the potentially confusing use of 'this' as an anaphoric reference to 'each' in the second sentence:

"Identify three advantages other than data independence of a DBMS's access of data over access of data in a conventional file via an access method. For each, explain how this is an advantage."

The instruction is probably more comprehensible if these potential problems are removed:

Data independence is one advantage of accessing data through a DBMS over accessing data in a conventional file via an access method. What are three other advantages?

.....

Another multi-purpose instruction word is 'determine'. Indeed in paper 25 alone it is used 5 times with varying meanings. In the first part of the following question, for example, it seems to mean 'find out and state', and in the second part 'list'. (The extract also suffers from faulty punctuation - a separate sentence should be used for the amplifying instruction 'using your

own words ..'.)

"(b) Given the four situations below, determine for each:

(i) if one or more array structures are needed ..

(ii) the steps necessary to implement the situations using your own words, not pseudocode."

.....

In the following extract the whole first sentence containing 'determine' could be replaced by the more straightforward instruction 'calculate':

"(a) Determine the following information.
(i) The averages of points scored by the home team and by the opponents .." (25)

.....

In the following two extracts the instructions exhibit different types of unnaturalness, which may cause the candidates a certain amount of confusion. For example:

"Pictorially describe the queue (using the same format as above) after each of the following operations." (23)

It is not immediately clear from this instruction what the candidate has to do. 'Pictorially describe' is a highly opaque way saying 'draw', and 'using the same format as above' is an unnecessarily inexplicit way to refer to a diagram. Candidates would need to spend less time working out the task if the instruction were worded as follows:

Draw a diagram like the one in Figure 1 above to show the queue after each of the following operations:

.....

"(a) Demonstrate in pictorial form similar to the diagram below all the intermediate states required to go from the initial state to the final state with three disks ... (27)

There are two potential problems here. Firstly, the instruction 'Demonstrate in pictorial form . . ' is an unnatural way of saying: 'Draw diagrams . . ' . Secondly, the candidate may be thrown off balance by the fact that reference is made to the (singular) 'diagram below' whereas in fact the question paper contains two diagrams, neither of which shows an intermediate state as required in the answer. A more immediately intelligible instruction would include a brief preamble,

and could read:

(a) The diagrams below show the initial state and the final state of a game to move 3 discs from one pole to another.

Draw a series of diagrams to show all the intermediate states that are needed to go from the initial state to the final state.

.....

Finally, it is worth pointing out the considerable lack of consistency across examination questions in the way candidates are instructed to write algorithms or draw flowcharts. The following imperatives are used in respect of algorithms: `supply, design, develop, produce, give, derive`. And 'flowchart' is preceded by; `construct, give, draw, draw up, prepare, design, write`. It is possible that candidates may, for example, think that a different type of response is required to the instruction '`Derive the algorithm ..`' than to the instruction '`Supply the algorithm ..`'. Although in general the candidate may experience only momentary uncertainty as to what is required, nevertheless there seems to be no good reason why these very common instructions should not be standardized using the most natural collocations: 'write an algorithm' and 'draw a flowchart'

5.11 Whole Question Analysis

So far, the many different types of potential difficulty have been identified and discussed largely in isolation, as exemplified in short text extracts. However, the comprehension of individual sentences or short passages may not be so difficult if the rest of the text is clear, and the organization of the question as a whole is logical and transparent. It is when a question has an accumulation of problems, even if these are seemingly trivial in themselves, that the candidate is likely to experience real difficulty. This is particu1ar1y so, of course, when the question contains a large amount of text, as in the case of the IB Computing Studies examination where a majority of the questions are longer than three quarters of an A4 page, and many are as long as 2 full A4 pages. The combination of the length of the question and the different difficulties it contains may cause the candidate to abandon his attempt to interpret exactly what is to be done and base his answer on a superficial understanding of the salient points. It is probably such behaviour which gives rise to the frequent complaint in the examiners' reports that candidates do not answer the question set.

It is important, therefore, to analyse one or more complete questions to show how the combination and accumulation of problems may considerably reduce the likelihood that the candidate will be able to understand the requirements of the question quickly and accurately.

5.11.1 Question 3 Paper M90/640/H(2) - (See Appendix 3a for the original text)

This first whole question chosen for detailed analysis places considerable demands on the candidate's reading ability. It is almost two pages long, it contains a large amount of uncommon lexis and it is about a topic of which few candidates are likely to have even indirect knowledge or experience. These facts alone are likely to make a quick and accurate understanding of the question tasks hard to

achieve, but the text is made more difficult because of many of the linguistic problems discussed earlier, and because it suffers from a certain lack of conceptual clarity.

The first paragraph introduces the theme and gives the necessary background information. Unfortunately, the meaning of difficult lexis such as 'infestation' and 'defoliation' is not made explicit; neither is the budworm clearly identified as an insect. Furthermore, the fact that 'tree' has been turned into a kind of mass noun (cf. the reference to 'balsam and spruce tree'), may cause momentary irritation to the attentive reader.

More serious is the fact that the stages/cycles of the 'competition' are not clearly designated as 'infested', then 'defoliated', then 'green'. Indeed, candidates have to work out that the sentence: "the spruce budworm defoliates balsam and spruce tree in one cycle" refers to the 'infestation' stage. This potential confusion is compounded in point 2 of paragraph 3, which lists the stages in a different order: "A site can be either green (g), infested (i), or defoliated (d)."

The situation described in the question may be a real one, but the reference to the three different kinds of tree, balsam, spruce and beech only lengthens the introduction and makes the candidate's task more difficult. This information could be omitted without prejudicing the nature of the tasks the candidate has to perform. And the same applies to the second paragraph, which apart from the first sentence contains information that is almost entirely superfluous.

A serious impediment to a clear understanding of the question as a whole arises in the third paragraph, which starts 'Assume the following rules: ..'. The problem lies in the conceptual ambiguity of the term 'cycle'. In the first paragraph, each of the three stages of the 'competition' process itself is designated as a cycle; the cycle of infestation being followed by the cycle of defoliation, which is in turn followed by the 'green' cycle, (the term given to the reafforestation stage. The first sentence of the second paragraph states explicitly: 'A forest may have many sites in different cycles". In other words, at any given time a site may be in the green cycle while its neighbouring sites may also be green, or defoliated, or infested. So the interpretation is that there are three cycles and that the term 'cycle' is synonymous with the stage or state of being infested, defoliated or green.

But in the third paragraph, 'cycle' comes to mean something different: Assumption 3. says that a green site remains green in the next cycle if it is not infested in the current cycle. And the example diagrams show the initial states of all 144 sites in the forest under the heading Cycle 0 and then the changes some time later under the heading Cycle 1. In other words, 'cycle' is no longer synonymous with the infested, defoliated or green state or stage of any individual site, but seems now to refer to an unspecified period of time during which some sites in the forest as a whole change but others do not. (And it is worth noting that neither of these senses of the word 'cycle' correspond to its everyday meaning, where it is regarded as a complete process consisting of a series of stages which recur in a regular sequence; for example, in the hydrological cycle water evaporates, rises into the air and later returns to the land as rain.) Clearly the candidate has a difficult task in sorting out this confusion before he can hope to interpret correctly what he has to do.

As far as the example is concerned, the two diagrams showing the changes in the state of the forest sites between Cycle 0 and Cycle 1 are intended to exemplify rules 1 to 6. It would probably help the candidate, however, if these changes were more immediately apparent, perhaps by being shown in bold type; and also if the fact that infested sites do not infest diagonally adjacent green sites were made more explicit.

There are two potential difficulties for candidates in question (a). Firstly, the task is not clearly marked off; it is set in two sentences in two paragraphs although in fact one only answer is required. Moreover, the instruction 'Supply your answer in a format similar to the earlier example indicating cycle 0 through cycle 2' is in itself problematic for various reasons:

1. 'Supply' is less natural than 'write'; and 'in a format similar to the earlier example' is less explicit than 'like the diagram in Figure 1 on page ..'.

2. The word 'indicating' is an amplifying instruction to the main task, but since there is no comma it could be erroneously interpreted as introducing an reduced relative clause referring to 'example'.

3. The Americanism 'cycle 0 through cycle 2' may not be clear to some candidates as meaning 'including cycle 2', particularly as the example which has just been referred to does not include cycle 2.

Secondly, slight confusion could be caused by the inconsistent use of the future tense and the tautology in the information that the array 'will contain the contents of the initial forest sites'. Strictly speaking, it is not the 'content' of the forest site which is stored in the array, but its state (green, infested, etc.)

Similar comments apply to Question (b), which otherwise has no significant difficulties. In the interests of consistency, however, it is preferable to replace 'Working with an initial forest' by the more conventional 'Assume an initial forest ,... "

In question (d) the difficulties for the candidate are in understanding the meaning of 'localised', and making the inference that removing the infestation 'naturally and safely' means 'without spraying'. In addition, the terms , 'embedded' , 'eradicate' and 'justifying' are likely to be troublesome. As far as the logic of the question is concerned, it seems superfluous to ask if the Forest Service can do anything natural about the infestation. Since the next question asks how long it will take to do so, the only possible answer is 'Yes'.

As for the way the last two tasks are formulated, the first can be expressed more naturally as: 'How many stages will it take? ' The second is badly punctuated and tautologous; the last three words can be omitted without loss.

The suggested revised version is at Appendix 3a.

5.11.2 Question 9 Paper M85/6.5/S - (See Appendix 4a)

This question will not be subjected to such a detailed analysis as the last one, although it contains areas of potential linguistic difficulty. It has been selected because the revised version serves as an example of how a clear signposting of the different parts of a question can help to alleviate the problems the candidate often faces in understanding what he has to do. The major difficulty with this particular question lies in the fact that the instructions in the last sentence are not explicit. The candidate reads through the whole text without knowing what he will have to do, and then is confronted by two commands which seem to belong to a different question - an impression that is

largely due to the use of the indefinite article in the bald instruction 'Prepare the flowchart for a program ..' . More immediately comprehensible would be the instruction: Draw a flowchart for the program that would operate the above system .. . It is suggested, however, that the candidate is better served if at the start of the question, he is given a clearer idea of the coming task(s). He is then able to read the remaining text in a more focused way.

Apart from these organizational problems, the following linguistic aspects of the question are worthy of brief comment. The points below refer to the part of the original question which is identified by the corresponding number - as shown in Appendix 4a:

1. The last two sentences of this paragraph are unnecessary to an understanding of the question, and are in any case likely to be part of most candidates' *drinks machine* schema.

2. The expression 'of each kind' is opaque in both sentences. Does it refer to the cups or the boxes?

3. The use of three different words to refer to the cups, i.e. *cups*, *containers* and *cartons*, is unnecessary and potentially confusing.

4. The embedding of 'on the computer' may cause the momentary interpretation that it is the form which is to be updated whereas it is in fact the stock levels that have to be updated, using the form.

The question has been radically reformulated to take account of the above comments. See Appendix 4b.

6. EXPERIMENT

The fundamental assumption underlying all the suggested revisions in section 5 of this paper is that the adapted versions would be more comprehensible to nonnative candidates than the original versions. This is easily claimed, but how can it be proved? Certainly, many changes involved the simplification of complex syntax or the replacement of an infrequent word by its more common equivalent. As such the texts in which those changes were made would probably score higher on a conventional test of readability such as the Flesch test. The value of such standardized measurements in this context, however, is highly dubious since they focus entirely on the linguistic features of the text and ignore the response of the reader, (see Carrell, 1987). Clearly, in order to prove the increased comprehensibility of the adapted texts we need to find a more responsive and reliable measurement than a readability formula.

If we return to our definition of comprehensible as meaning *'capable of being understood quickly'*, it is clear that in order to measure the comprehensibility of examination questions, in their original or adapted form, we must measure the candidate's understanding of the task he has to perform and measure the speed at which he has arrived at that understanding. Measuring speed is not a problem, but measuring understanding certainly is. Understanding is a cognitive process that cannot be examined directly, so we have to measure it indirectly by having the reader produce evidence of his understanding. In the case of an examination question, of course, one type of evidence of understanding is the candidate's written response to the question task, and it might seem that this would be the best object of an investigation into comprehension. But as we noted earlier, although a full and accurate answer would in most cases be sufficient evidence that the candidate has understood the question, it is clearly not admissible to claim, conversely, that an incomplete or wrong answer is proof of the candidate having failed to understand the text. The nature of the answer given tells us

nothing about whether or not the question was comprehensible in the sense that the candidate was able to understand it quickly.

Probably the best way that the comprehensibility of an examination text could be measured would be to allow the candidate as long as he wanted to understand the task(s) and then have him relate to the writer of the question (i.e. the examiner) exactly what he thinks he has to do. If this were done with many candidates, then any questions which required a disproportionately long processing time, and any questions which tended to reveal a discrepancy between the examiner's and the candidate's understanding of the task to be performed, could be regarded as having low comprehensibility.

There is another, easier way of measuring comprehensibility. It is clear from the previous discussion that comprehensibility is not an intrinsic quality of a text, but depends on the reader's response to that text, as influenced by numerous variables such as his language competence, background knowledge and reading purpose. It seems reasonable to suggest, therefore, that comprehensibility could also be measured in terms of the reader's subjective assessment of how easy or difficult it is to make sense of a text. In other words, if he thinks a text is difficult, it is difficult. Such a subjective assessment may have direct consequences in an examination, since candidates who think that the language of a particular question is difficult are more likely to believe that they have not fully understood. They may either avoid doing the question in the first place or will be more tentative in their answer. (For a further discussion of the mentalist approach to assessment of reading difficulty, see Cohen, 1984)

On the basis of these suggestions as to how the comprehensibility of examination questions can be measured, an experiment was designed to test, among other things, the hypothesis that the adaptation of examination texts can result in increased comprehensibility. The experiment was to be carried out with Computing Studies students at Frankfurt International School, and the intention was to compare the students' responses to original and adapted texts in two identical tests to be conducted at an interval of about one month. Unfortunately, for various reasons, only 8 students participated in the experiment, and of those only 2 completed both tests. Furthermore, none of the participants was a nonnative speaker who was still in the school's ESL programme - the group most likely to experience problems with the language of the question texts, and most likely to benefit from adaptations. As a result, neither the main hypothesis nor the subsidiary ones (see below), could be tested. Nevertheless, some interesting insights were gained, which will be of use if the project is to be repeated in the future. The following is a description of the experiment as it took place - for a more detailed account of how the complete experiment was to be carried out, see Appendix 5 (not availabe online).

PARTICIPANTS

The eight students participating in the experiment were all in their final year at Frankfurt International School and candidates for the May 1993 IB Computing Studies examination. Two of them were native English speakers; the other 6 were nonnative speakers but all of these were proficient enough in English to have left the school's English as a Second Language (ESL) programme.

METHOD

Twelve examination questions were chosen as exemplifying the problems discussed in the previous sections of this analysis. If the question was short, it was used in its entirety, otherwise only the first part of it was used. An adapted version of each question (or question part) was produced, which was checked by a Computing Studies teacher to verify that the changes had not altered the meaning of text or the nature of the task to be performed.

Half of the original versions and half of the adapted versions (i.e. a total of 12 different questions) were presented to the students one at a time on a computer screen, using a specially designed computer program (Appendix 6 - not available online). Each extract remained on the screen for as long as the student wanted. When he thought he had understood the contents, he pressed a key and his time was recorded. The text then disappeared and he was asked to rate its difficulty on a scale from 1 (easy) to 5 (difficult). Having done this, he was presented with a true/false statement about the extract to test his comprehension. (This order of presentation, i.e. text - rating - question, was to avoid the possibility that the student would rate the difficulty of the question rather than the difficulty of the text.) To discourage guessing, the student was allowed to chose a *'don't know'* option if he really felt unsure of the correct answer.

As noted above, there are many problems associated with measuring understanding, and clearly a true/false test cannot hope to do this validly or reliably. However, the primary aim of the true/false statements was not in fact to check comprehension. The intention was to give the students some kind of reading purpose. It was felt that they would not necessarily read carefully enough if no task were to follow.

FINDINGS

Appendix 7 (not available online) gives a comparison of the number of words and the Flesch ratings of the original and simplified versions, together with a summary of the results of the experiment as it took place. Appendix 8 (not available online)contains the original and adapted versions of each of the twelve chosen texts, together with a summary of the responses to the particular version used in the first test. As far as the adaptations are concerned, an interesting point to make is that although the changes in all cases except one resulted in an increase in readability as measured by the Flesch index, only half of the 12 adapted texts were shorter than the original versions. The length of questions 1, 5, 7, 9, 10 and 11 was increased by the adaptations.

At first glance, there is some slight evidence that the adaptations increased the comprehensibility of the texts. For example, the adapted texts were rated on average 0.6 of a point easier than the original versions. The mean time taken to read the adapted versions was about 5 seconds shorter than that needed for the original versions, even though the adaptations hardly changed the mean length of the texts (from an average of 94.3 words per original text to 95.0 words per adapted text). It should be noted, however, that the number of words in the individual passages correlated quite highly with the difficulty rating that passage was awarded ($r_{observed}$ = .59, Pearson). And that likewise there was a reasonably high correlation between the number of words in a text and the time taken to read it ($r_{observed}$ = .61).

The true/false questions about the adapted texts were answered correctly by an average of 1 more student per question than the original texts, but this is balanced by the fact that more students per question got the questions about the adapted texts wrong. The apparent illogicality is explained by the fact that an average of 1.7 more students per question did not feel able to give an answer to the true/false statement about the original texts. Of course, no conclusions should be drawn from these last data because there is no way of knowing whether the results had anything to do with the comprehensibility of the texts. It could be, for example, that more students answered the questions about the adapted texts correctly because the questions themselves were simpler or more comprehensible than the questions about the original versions.

Furthermore, the findings of parts of the experiment are unreliable since it was clear that one or two students had quickly given up on some of the longer texts, and had guessed at the answer to the corresponding true/false statement.

DISCUSSION

As Davison & Kantor (1982) state: "*The effects of specific changes made in adaptation are very hard to test experimentally*." Clearly it would be very unwise to attempt to draw any conclusions from a half-completed experiment with unreliable data. Perhaps the most we can do is to note that, at the very least, the adapted IB Computing Studies examination texts did not have a lower comprehensibility than the original versions, as measured by the time it took to decode them, and by the difficulty ratings they were given. Clearly, if the experiment is to be undertaken again in the future, it is imperative that both tests should be taken so that a direct comparison of the responses to the adapted and original versions is possible. It is also essential that nonnative candidates who are still in the ESL programme should participate in the experiment.

Furthermore, certain alterations need to be made in order to overcome some of the problems that have emerged from the first trial run. For example, a way needs to be found to ensure that participants treat the tests seriously. It has been noted how some students obviously did not attempt to read the some of the longer texts carefully, and another student, whose answers were not taken into account in the findings discussed above or shown in the appendix, rated all the texts as easy. A further problem is that there is no way of knowing to what extent the difficulty rating given reflected the nature of the task or the concepts contained in the question/question extract, rather than the difficulty of the language in which those concepts or that task was expressed.

There are no easy ways to overcome these problems other than to give the participants very explicit instructions and to appeal to their conscientiousness. It would also be helpful to allow them more than enough time to finish the test; and the promise of a prize for the most correct answers could also prove an effective incentive. At the analysis stage the answers should be carefully sifted to remove the obvious problem cases.

Consideration needs to be given also to the use of the true/false comprehension questions. Their inclusion in the experiment was simply to provide the students with a reading purpose, but it would be useful, nevertheless, to have a more sophisticated and reliable way of testing comprehension of the text. On the assumption that an individual interview with each candidate about each question is not feasible, then the best method would seem to be to administer a multiple-choice test. However, the difficulties of producing an effective and reliable one should not be underestimated (Thynne, 1974).

If an experiment can be designed that overcomes the difficulties discussed above, then it should be possible to test the following hypotheses:

1. that all students (native and nonnative speakers) will rate the simplified texts as easier than the original ones;

2. that nonnative ESL students will give (all) texts a higher difficulty rating than native speakers;

3. that all students will need less time to decode the simplified texts than the original ones;

4. that nonnative ESL students will need more time to decode both types of text than native speakers;

5. that all students will answer more questions about adapted texts correctly than about original texts.

DISCUSSION

The analysis of the language of the IB Computing Studies questions and the experiment that followed it have important implications for ESL teachers preparing nonnative candidates for the examination. Clearly, the ESL teacher not only has a general responsibility to improve the student's overall English proficiency - i.e. to help him reach a threshold where his reading skills are not likely to be short-circuited by his lack of language competence. The teacher can also give more directed instruction to help the student prepare for the examination.

An important part of this instruction would be the explicit teaching of the skills of reading. Although L2 texts have always featured in language teaching, they have often been used primarily as a source of new vocabulary or as a means of exemplifying target language forms, (Johns & Davies, 1983). This may or may not be justified depending on the teaching situation, but students also need explicit training in the selection and use of the most appropriate reading strategies to be employed in reading a particular text for a particular purpose. In the case of an examination, for example, they need instruction and practice in how to achieve a full and exact understanding of what the question requires them to do in as little a time as possible. This is not the occasion to discuss in detail how such skills can be taught (see Nuttall, 1982, and Williams, 1983, for good practical advice in this area), but it is worth emphasising that it presupposes working with target texts, i.e. examples of the examination questions themselves.

Another important reason for working with target texts is that it gives the students a chance to familiarize themselves with the typical organization of such texts. In other words, in the case of the IB Computing Studies examination, it helps them develop a formal schema for the way questions in general are structured, that will help them in their efforts to understand individual questions. Barnett (1989) stresses the importance of formal schemata:

> *They are important because they define reader expectations about how pieces of textual information will relate to each other and in what order details will appear.*

Both Carrell (1988) and Grabe (1991) have shown that explicit training in recognizing and analysing the rhetorical organization of text can facilitate reading comprehension. In view of the importance for candidates of developing a knowledge of the typical structure of examination texts, i.e. of acquiring a formal schema, it would not seem unreasonable to suggest that the examiner has a responsibility to ensure that questions are structured in a consistent and logical fashion within and across examination papers.

One final crucial area in which the teacher can help prepare students for the examination is by training them in the proper use of a dictionary. That nonnative candidates are allowed to use dictionaries in the examination is to some extent a dubious advantage. Bensoussan & Sim (1984), for example, found that many university students were inefficient and ineffective users of dictionaries in tests. Predictably, students who did not use a dictionary finished faster than those who did, but rather less predictably, there were only insignificant differences in the scores of the users and non-users. It would be tempting to conclude from this study that students should be encouraged not to consult their dictionaries at all, but this would seem to ignore the realities of a situation in which students rightly feel that they need a complete and full understanding of the text they are reading. Of course, it is highly debatable whether an adequate comprehension of an examination question depends on understanding every word, but what is at issue here is the candidate's response to unknown

vocabulary. Eskey and Grabe (1988) comment on a typical response and its consequences:

> *Second language readers are frequently panicked by the occurrence of unknown words, and they stop reading to look them up in dictionaries, thereby interrupting the normal reading process.*

On the assumption that nonnative candidates will typically use a dictionary in the examination whenever they encounter a difficult word - an assumption that needs to be established empirically - the teacher has two responsibilities. Firstly, he must ensure that the student's dictionary is a good one, and secondly, he must ensure that the student knows how to use it quickly and accurately. This is no easy matter, as Gremmo (1985) points out:

> *To use a dictionary properly, the reader should analyse the syntactic and semantic context of the unknown word before he even opens it (sic). In other words he should already have some idea about what it is he is looking for, because it is only this analysis which is going to enable him to choose the entry and the sense which correspond to the passage in question.*

This type of training will clearly require a course of instruction and ideally will again entail working part of the time with target texts, i.e. with past examination papers. This, in turn, presupposes close cooperation with the specialist teacher. This is a demanding exercise, but is clearly an integral part of the responsibility of the ESL teacher working in an international school,

However, it needs to be emphasized; that the examiner has a responsibility in this area too. It is not unreasonable to recommend, in view of the predictable difficulties caused by unknown words and the problems of decoding them efficiently using a dictionary, that the examiner should as a rule choose the more frequent word ahead of the less frequent one in cases where the meaning would otherwise be unaffected.

8. CONCLUSION

Examination questions are a very special kind of text. They are read by students under considerable pressure - the general pressure of knowing that the following three hours could have a significant effect on their future lives, and the more specific pressure of trying to balance the amount of time needed to understand the instructions in order to answer exactly the question set with the amount of time needed to give a full and correct answer.

While an examination board clearly cannot be taken to task if questions are found difficult purely by virtue of the fact that candidates are lacking in basic English proficiency, it does bear full responsibility for ensuring that the language of the examination texts conveys the instructions and any necessary explanatory information concisely and unambiguously, particularly when that examination is expressly targetted at nonnative candidates. For as soon as a subject examination becomes a test of English it starts to lose its validity as test of knowledge about that subject.

The preceding analysis has attempted to show that the International Baccalaureate Computing Studies examination is to a significant degree a test of English.

BIBLIOGRAPHY

ALDERSON, J and URQUHART, A (1985) What is Reading? In: ALDERSON, J and URQUHART, A *(eds.) Reading In a Foreign Language* London Longman

BARNETT, M (1989) *Reading in a Foreign Language: More than meets the eye* Rowley Newbury House

BATEN, L and CORNU, A (1984) Reading strategies for LSP text In: PUGH, A and ULIJN, J (eds.) *Reading for Professional Purposes* London Heinemann

BEAUMONT, C (1982) Reading Relative Clauses *Journal of Research in Reading* 5 (1)

BENSOUSSAN, M and SIM, D (1984) The Effect of Dictionary Use on EFL Test Performance *Reading In a Foreign Language* 2(2)

BERMAN, R (1984) Syntactic components of the foreign language reading process In: ALDERSON, J and URQUHART, A (eds.) *Reading in a Foreign Language* London Longman

BLAU,E (1982) The Effect of Syntax on Readability for ESL Students in Puerto Rico *TESOL Quarterly* 16(4)

BROWN, R (1987) A comparison of the comprehensibility of modified and unmodified reading materials for ESL *University of Hawaii Working Papers in ESL*

CANDLIN, C (1979) Discourse analysis: interpretive strategies and the process of reading Presentation at a regional university teachers of English in Israel, Jerusalem

CARRELL, P (1982) Cohesion is not coherence *TESOL Quarterly* 16(4)

CARRELL, P (1987) Readability in ESL *Reading In a Foreign Language 4*

CARRELL, P (1988a) Introduction: Interactive approaches to second language reading In: CARRELL, P., DEVIME. J. and ESKEY, D (eds.) Interactive *Approaches to Second Language Reading* Cambridge Cambridge University Press

CARRELL, P (1988b) Interactive text processing: implications for ESL/second language reading classrooms In: CARRELL, P., DEVINE, J. and ESKEY, D (eds.) Interactive *Approaches to Second Language Reading* Cambridge Cambridge University Press

CASSELS, J and JOHNSTONE, A (1978) Language in Chemistry .Report *of University of Glasgow Science Education Research Group*

CLARKE, M and SILBERSTEIN, S (1975) Toward a Realization of Psycholinguistic Principles In: MACKAY, R (et al, eds.) *Reading In a Second Language* Rowley Newbury House

CLARKE, M (1988) The short-circuit hypothesis of ESL reading -or when language competence interferes with reading performance In: CARRELL, P., DEVINE, J. and ESKEY. D (eds.) Interactive *Approaches to Second Language Reading* Cambridge Cambridge University Press

COHEN, A (1984) The use of mentalist measures in determining LSP reading problems In: PUSH, A and ULIJN, J (eds.) *Reading for Professional Purposes* London Heinemann

COHEN, A (1990) *Language Learning* New York Newbury House

COHEN, A (et al.) (1979) Reading English for Specialized Purposes: Discourse Analysis and the Use of Student Informants *TESOL Quarterly* 13(2)

COOPER, M (1985) Linguistic competence of practised and unpractised non-native readers of English In; ALDERSON, J and URQUHART, A (eds.) *Reading in a Foreign Language* London Longman

COWAN, J (1974) Lexical and Syntactic research for the Design of EFL reading materials *TESOL Quarterly* 8(4)

CUMMINS, J (1991) Language Development and Academic Learning In: MALAVE, L & DUQUETTE, G (eds.) *Language, Culture, Cognition* Clevedon Multilingual Matters

CUMMINS, J (1984) *Bilingualism and Special education: Issues in Assessment and Pedagogy* San Diego College Hill

DAVIES, A (1985) Simple, simplified and simplification: what is authentic? In: ALDERSON, J and URQUHART, A (eds.) *Reading in a Foreign Language* London Longman

ESKEY, D and GRABE, W (1988) Interactive models for second language reading In; CARRELL, P, DEVINE, J and ESKEY, D (eds.) Interactive *Approaches to Second Language Reading* Cambridge Cambridge University Press

EVANS, E (1988) Advanced ESL Reading - Language Competence Revisited *System* 16(3)

FODOR, J and GARRET, B (1967) Some syntactic determinants of sentential complexity *Perception and Psycholinguistics* 2(7)

FRANCIS, G (1987) Anaphoric Nouns JELJR Journal *11* Centre for English Language Studies University of Birmingham

FULCHER, G (1989) Cohesion and Coherence in Theory and Reading Research Journal of *Research in Reading 12*

GOODMAN, K (1967) Reading: a psycholinguistic guessing game Journal or the *Reading Specialist* 6(1)

GRABE, W (1988) Reassessing the -term "interactive" In:
CARRELL, P, DEVINE, J and ESKEY, D (eds.) Interactive *Approaches to Second Language Reading* Cambridge Cambridge University Press

GRABE, W (1991) Current Developments in Second Language Reading Research *TESOL Quarterly* 25(3)

GREMMO, M (1985) Learning a language - or learning to read? In: RILEY, P (ed.) *Discourse and Learning* Harlow Longman

HALLIDAY. M and HASAN. R (1976) *Cohesion In English* London Longman

HARRISON, C (1980) Readability in the *Classroom* Cambridge Cambridge University Press

HONEYFIELD, J (1977) Simplification *TESOL Quarterly* 11(4)

HUDSON, T (1988) The effects of induced schemata on the "short circuit" in L2 reading In:
CARRELL, P., DEVINE, J. and ESKEY, D (eds.) Interactive *Approaches to Second Language Reading* Cambridge Cambridge University Press

JOHNS, T (1980) The Text and its Message: an Approach to the Teaching of Reading Strategies for Students of Development Administration In: FABER, H and MALEY, A (eds.) *Leseverstehen 1m Fremdsprachenunterrlcht* Munich Goethe Institut

JOHNS, T and DAVIES, F (1983) Text as a Vehicle for Information: the Classroom Use of Written Texts in Teaching Reading in a Foreign Language *Reading in a Foreign Language 1(1)*

JOHNS, T and DUDLEY-EVANS, T (1988) An experiment in the team-teaching of overseas

students of transportation and plant biology In: SWALES, J *Episodes In ESP* Hemel Hempstead Prentice Hall

JOHNSON, P (1987) language complexity *Uni*versity of *Hawaii*
Effects on reading comprehension of and cultural background of a text *Working Papers in ESL*

JOLLY, D (1978) The establishment of a self-access scheme for intensive reading Paper presented at Goethe Institut, British Council Collegium on Reading, Paris

KAHN, J (ed.) (1985) *The* London Reader's Digest *Right Word at the Right Time*

KLARE, G *Quarterly* (1974) Vol X Assessing Readability .Reading- *Research*

LAUFER, B (1992) Reading in a foreign language: how does L2 lexical knowledge interact with the reader's general academic ability? *Journal of Research In Reading* 15(2)

LONG, M, (et al.) (1991) Simplification or Elaboration? The effects of two types of text modifications on foreign language reading comprehension *University of Hawaii Working Papers In ESL*

MACKAY, R, BARKMAN, B and JORDAN, R (1979) (eds.) *Reading in a Second Language* Rowley Newbury House

MCCARTHY, M (1988) Questions of Naturalness in Language In: MCCARTHY, M (ed.) *Naturalness In Language* ELR Journal Vol 2 Birmingham University

MEARA, P (1984) Word recognition in foreign languages In: PUGH, A and ULIJN, J (eds.) *Reading for Professional Purposes* London Heinemann

NUTTALL, C (1982) *Teaching Reading Skills In a Foreign Language* Oxford Practical Language Teaching

O'DONNELL, W (1968) An Investigation into the Role of Language in a Physics Examination Monograph Moray House Edinburgh

PARKER, K and CHAUDRON. C (1987) The Effects of Linguistic Simplification and Elaborative Modifications on L2 Comprehension *University of Hawaii Working Papers In ESL*

PERERA, K (1982) The Assessment of Linguistic Difficulty in Reading Material In: CARTER, R (ed.) *Linguistics and the Teacher* London Rout ledge and Kegan Paul

PERKINS, K (et al.) (1989) First and Second Language Reading Comprehension *RELC Journal* 20(2) Singapore SEAMEO Regional Language Centre

QUIRK, R (et al.) (1985) *A Comprehensive Grammar of the English Language* London Longman

RUMELHART, D (1977) Toward an interactive model of reading In: DORNIC, S (ed.) Attention and *performance VI* New Jersey Er1baum

SAMUELS, S and KAMIL, M (1988) Models of the reading process In: CARRELL, P., DEVINE. J. and ESKEY, D (eds.) Interactive *Approaches to Second Language Reading* Cambridge Cambridge University Press

SINCLAIR, J (ed.) (1987) *Collins Cobuild English Language Dictionary* London Collins

SMITH, F (1978) *Reading* Cambridge Cambridge University Press

STANOVICH, K (1980) Toward an interactive-compensatory model of individual differences in the

development of reading fluency *Reading Research Quarterly 16*

SWAFFAR, J (1991) *Reading for Meaning* Hemel Hempstead Prentice Hall

SWALES, J (1982) Examining Examination Papers *English Language Research Journal 3*

THYNNE, J (1974) *Principles of Examining* University of London Press

ULIJN, J and STROTHER, J (1990) The effect of syntactic simplification on reading EST texts in LI and L2. *Journal of Research in Reading* 13(1)

URQUHART, A (1984) The effect of rhetorical ordering on readability In: ALDERSON, J and URQUHART, A (eds.) *Reading in a Foreign Language* London Longman

WEBER, R (1984) Reading *Annual Review of Applied Linguistics*

WIDDOWSON, H (1978) *Teaching Language as Communication* Oxford Oxford University Press

WIDDOWSON, H. (1990) *Aspects of Language Teaching* Oxford Oxford University Press

WILLIAMS, R (1983) Teaching the Recognition of Cohesive Ties in Reading in a Foreign Language *Reading in a Foreign Language* 1(1)

WILLIAMS, R and DALLAS, D (1984) Aspects of vocabulary in the readability of content area L2 educational textbooks: a case study In: ALDERSON, J and URQUHART, A (eds.) *Reading in a Foreign Language* London Longman

WILLIAMS, R (1984) A cognitive approach to English nominal compounds In: PUGH, A and ULIJN, J (eds.) *Reading for Professional Purposes* London Heinemann

YORIO, C (1971) Some sources of reading problems for foreign language learners *Language Learning 21*

APPENDICES

1. Statement about the International Baccalaureate programme - not available online

2. List of Examination Papers - not available online

3a. Original Question (M90/640/H(2)) - see section 5.11.1

3b. Revised Question (M90/640/H(2))

4a. Original Question (M85/6.5/S) - see section 5.11.2

4b. Revised Question (M85/6.5/S)

5. Letter to the computing studies teachers responsible for administering the tests - see section 6 - not available online

6. Computer Program used in experiment - not available online

7. Summary sheet of experiment data - not available online

8. Original/adapted versions of texts used in experiment plus additional data - not available online

APPENDIX 3a - ORIGINAL QUESTION (Section 5.11.1)

This question involves the theme of "competition between species through infestation". As an explanation of this theme the spruce budworm defoliates balsam and spruce tree in one cycle. The defoliated site then grows beech trees in the next cycle which are ignored by the budworm. Eventually in the next cycle, the beech trees are replaced by balsam and spruce trees and the competition begins again.

A forest may have many sites in different cycles. The Forest Service can spray sites If they are infested but usually will not since the spray is unnatural and damaging. Occasionally, a natural and safe approach for defoliation can be introduced for control.

Assuming the following rules:

1. A forest is composed of square sites.

2. A site can be either green (g), infested (i), or defoliated (d).

3. A green site (g) remains green in the next cycle if it is not infested during the current cycle (see rule 5).

4. A defoliated site (d) becomes a green site (g) in the next cycle.

5. An infested site (i) infests its neighbours to the North, South, East and West only if they are currently green sites (g).

6. An infested site (i) becomes a defoliated one (d) in the next cycle.

7. An infested site (i) cannot infest beyond the boundaries of the forest.

Below is an example of an initial forest (12 sites by 12 sites) at cycle 0 and the results of the forest at cycle 1.

CYCLE 0

```
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
ggggiiiigggg
ggggddddgggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
```

CYCLE 1

```
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
ggggiiiigggg
gggidddidggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
```

Answer the following questions.

(a) A two-dimensional array of single character attributes ('g', 'i', 'd') called FOREST will contain the contents of the initial forest sites. What are the FOREST (6 sites by 6 sites) characteristics at cycle 2 given the initial FOREST contents (cycle 0) of:

```
FOREST  (1,3)  <--  'i'  FOREST  (2,6)  <--  'i'  FOREST  (4,3)  <--  'i'
FOREST  (2,3)   <-- 'd'   FOREST (3,3)    <-- 'd'   FOREST (5,3)   <-- 'i'
```

All other sites are green ('g'). Supply your answer in a format similar to the earlier example indicating cycle 0 through cycle 2.

(b) Working with an initial forest of size 12 sites by 12 sites, write the algorithm (not the program language code) that moves the forest from one cycle to another. A two-dimensional array of single character attributes ('g', 'i', 'd') called FOREST will contain the contents of the initial forest sites. You may wish to utilise a similar array, NEWFOREST, to hold the new site characteristics as the cycle progresses.

(c) The Forest Service will spray the forest at the end of any cycle if more than *50%* of the sites are nongreen. Write a subalgorithm that determines the percentage of nongreen sites and generates the message to spray, if necessary, at the end of the cycle.

(d) If the infestation is localised (a single 'i' site or all 'i' sites are neighbour (N.S.E.W) of at least one other 'i' site with no embedded green sites), can the Forest service (sic) do something to eradicate the infestation naturally and safely? How long will it take in terms of cycles? Explain your answer justifying your conclusion.

APPENDIX 3b

REVISED QUESTION (Section 5.11.1)

SITUATION:

This question is about the effects of an insect called the budworm on trees in a forest. Some parts of the forest are infested by the budworm, which eats all the leaves on the trees. The defoliated trees die but after a time new trees begin to grow in their place and the site becomes green again. Sooner or later the new trees are themselves infested by the budworm and the whole process starts again.

ASSUMPTIONS:

1. The forest is composed of many square sites.

2. A stage is a period of time in which some forest sites are infested (i), others are defoliated (d), and the rest are green (g).

3. An infested site (i) always becomes defoliated (d) in the next stage.

4. A defoliated site (d) always becomes green (g) in the next stage.

5. i.) If a green site (g) is next to an infested site (i), it will become infested in the next stage.

ii.) If a green site (g) is **not** next to an infested site (i), it will stay green in the next stage.

6. Green sites do not become infested if they are *diagonally* next to infested sites.

EXAMPLE:

Figure 1 below is an example of the changes in a forest (12 sites by 12 sites) from stage 1 to stage 2.

Figure 1:

STAGE 1

gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
ggggiiiiggggg
ggggddddggggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg

STAGE 2

gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
gggg**iiii**gggg
ggg**iddddi**ggg
gggg**gggg**gggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg
gggggggggggg

Note: The sites in bold are the only ones that have changed from stage 1 to stage 2.

TASKS:

Task 1: A forest contains 36 sites - (6 sites by 6 sites). Draw three diagrams like the ones in figure 1 to show the forest at stage 1, stage 2 and stage 3.

The following information is given:

There is a two-dimensional array of single character attributes ('g', 'i', 'd') called FOREST" which has the following elements at stage 1:

FOREST (1,3) <-- 'i' FOREST (2,6) <-- 'i' FOREST (4,3) <-- 'i'
FOREST (2,3) <-- 'd' FOREST (3,3) <-- 'd' FOREST (5,3) <-- 'i'

All the other elements of the array are 'g' (green).

Task 2: A second forest contains 144 sites (12 sites by 12 sites). Write the **algorithm** that moves the forest from one stage to the next. (Do not write program language code.)

The following information is given:

> The states of the sites at stage 1 are held in a two-dimensional array of single character attributes ('g', '1', 'd') called FOREST.

(HINT: You may wish to use an array "NEWFOREST" to hold the interim characteristics of each site.)

Task 3: The Forest Service sprays the forest with chemicals at the end of a stage if more than 50% of the sites are 'nongreen'. Write an algorithm that calculates the percentage of the 'nongreen' sites and generates the message to spray, if necessary, at the end of the stage.

Task 4: Explain how the Forest Service could stop the infestation without spraying. How many stages would it take?

Assume that the infestation in a forest is localised. (This means that there is only one 'i' site" or that all the 'i' sites are next to at least one other 'i' site and do not completely surround any green sites.)

APPENDIX 4a ORIGINAL QUESTION (Section 5.11.2)

A school uses a vending machine to provide hot drinks for students.

Drinks dispensed are:

a) coffee
b) tea
c) soup
d) chocolate

The machine holds the cups, which are supplied complete with the ingredients for making each drink. A maximum of 60 cups of each drink can be stored in the machine when it is refilled. An unlimited supply of hot water is available. The machine, on the insertion of the correct money. dispenses a cup of the chosen drink readymade with the hot water[1].

The cups are supplied in boxes of 300 of the same kind[2]. The school store-room holds at most, 5 boxes of the same kind[2]. The number of cups[3] of each type in the machine is checked at the end of each day. More are added if the level of any particular drink has dropped below 30.

The school microcomputer is to be used to initiate re-ordering of boxes of containers[3] as and when these are required. Every morning the person who refills the machine each day provides the school office with a form to update, on the computer[4], the stock levels. The computer, as part of the updating process, issues an appropriate order to the company supplying the boxes of cartons[3] when the number of boxes of any type of drink falls below two.

Prepare the flowchart for a program and design a form for use by the person responsible for filling the machine.

APPENDIX 4b REVISED QUESTION (Section 5.11.2)

SITUATION:

Imagine that you have been asked to help computerize the system for ordering cups for your school's drinks machine, which sells coffee, tea, soup and chocolate.

This is how the system is to operate:

> The drinks machine is checked at the end of each school day. If there are less than 30 cups of any drink, more cups are added. The school office is then informed how many drinks of each kind were sold that day. When there are not many boxes of cups left, the drinks machine company is contacted and asked to send more boxes.

ASSUMPTIONS:

1. The maximum number of cups that the machine can hold for each drink is 60.

2. The cups come in boxes of 300. Each box contains only one type of drink.

3. The school can store a maximum of 5 boxes of each drink.

4. An order is sent to the drinks company whenever the number of boxes of any type of drink falls below 2.

TASKS:

(a) Draw the flowchart for a program which would issue the order for new boxes of cups when they are needed.

(b) Design a form that is filled in by the person who checks the machine each day and used by the school office to order more boxes of cups when needed.