

Research Notes

Contents

Editorial Notes	1
A cognitive processing approach towards defining reading comprehension	2
Applying a cognitive processing model to Main Suite Reading papers	11
A corpus-informed study of specificity in Financial English: the case of ICFE Reading	16
Exploring lexical differences in General English Reading papers	22
Text organisation features in an FCE Reading gapped sentence task	26
Conference reports	31
Recent publications of interest	34
IELTS award news	35

Editorial Notes

Welcome to issue 31 of *Research Notes*, our quarterly publication reporting on matters relating to research, test development and validation within Cambridge ESOL.

In this issue we focus on the skill of reading, a component in all of our language assessments and teaching awards. We approach reading in a number of ways: from the general to the specific: from the theoretical (defining the construct); through the practical (operationalising the construct) to corpus-informed studies of reading vocabulary across the proficiency continuum and finally to the thematic organisation of reading passages.

In the opening article, Cyril Weir and Hanan Khalifa describe the mental processes readers use to comprehend reading texts as a means of defining Cambridge ESOL's construct of reading that our language assessments purport to test. In the following article, Cyril Weir and Hanan Khalifa apply this cognitive processing approach to defining reading comprehension to the Cambridge ESOL Main Suite examinations in English, focusing on two levels. This analysis of two of our Main Suite examinations – the Preliminary English Test (PET) and the First Certificate in English (FCE) – both which have a long pedigree and the support of considerable expertise and experience in pedagogy, is an attempt to ground our theory in practice.

The following three articles present corpus-informed studies of Reading texts in different types of English test. Firstly, Angela Wright investigates the specificity of Financial English reading texts compared to Business English and General English reading texts, using the new International Certificate in Financial English (ICFE) Reading paper as a case study. Fiona Barker then explores lexical differences more broadly in General English Reading papers, comparing reading passages at five Common European Framework of Reference (CEFR) levels to inform criterial differences between proficiency levels, using Main Suite Reading papers. In the third study, Glyn Hughes compares the text organisational features of reading passages from a First Certificate in English (FCE) paper with the original source text, seeking evidence for how candidates interact with reading passages and implications for training materials writers.

We then report on recent conferences and courses attended by Cambridge ESOL staff, including several ALTE and IELTS related events, and review recent publications of interest. Finally, we include the latest IELTS award news which consists of announcements of the recipients of the latest IELTS Joint-funded Research Programme and IELTS Masters Award together with calls for submissions for the next round of both schemes.

Cambridge Assessment is celebrating 150 years throughout 2008 so why not join us at the ALTE 2008 and IAEA 2008 conferences we are hosting to help mark the occasion.

Editorial team for Issue 31: Fiona Barker, Hanan Khalifa and Kirsty Sylvester.

A cognitive processing approach towards defining reading comprehension

CYRIL WEIR UNIVERSITY OF BEDFORDSHIRE
HANAN KHALIFA RESEARCH AND VALIDATION GROUP

Introduction

In this article we focus on a *cognitive processing* approach as a theoretical basis for evaluating the cognitive validity of reading tests.¹ This approach is concerned with the mental processes readers actually use in comprehending texts when engaging in different types of real-life reading. However, we first start by a brief review of other approaches that attempted to establish what reading comprehension really involves.

A factorial approach to defining reading comprehension

From the 1960s onwards there has been a strong interest in the issue of the divisibility of reading for testing purposes. In pursuit of this divisibility hypothesis-testing researchers often adopted a purely quantitative approach to establishing what reading is by a *post hoc*, factorial analysis of candidate performances in reading tests. This methodology tells us whether the different reading items we have included in our tests load on the same factor.

Davis (1968) provides an early example of empirical research into the factors contributing to successful test performance. He employed eight subtests designed to measure distinct operations. When applying factor analysis, five factors showed appreciable percentages of unique variance and were consistent across the test forms which made him argue that ‘comprehension among mature readers is not a unitary mental operation’ (Davis 1968:542). The factors were: recalling word meanings, drawing inferences, recognising a writer’s purpose/attitude/tone, finding answers to explicit questions and following the structure of a passage. Using factor analysis on engineers’ reading performance, Guthrie and Kirsch (1987) identified two factors. Firstly, reading to comprehend, which involves reading carefully to understand the explicitly stated ideas, was clearly differentiated from tasks involving reading to locate information which required selective sampling of text (see Weir et al. 2000 for similar findings).

However, these findings in favour of divisibility of the reading construct are not shared by other researchers. Rosenshine’s (1980) review of factor analytic empirical studies suggests that different analyses yielded different unique skills. Even though some skills emerged as separate, the results were not consistent across the studies which led him to conclude by saying ‘at this point, there is simply no clear evidence to support the naming of discrete

skills in reading comprehension’ (Rosenhine 1980:552). Schedl et al. (1996) looked at the dimensionality of TOEFL reading items specifically in relation to “reasoning” (analogy, extrapolation, organisation and logic and author’s purpose/attitude) as against other types (primarily items testing vocabulary, syntax and explicitly stated information). Their study did not support the hypothesis that the “reasoning” items measured a separable ability factor.

Limitations of the factorial approach

The factorial approach focuses on the separability of the capabilities that a reader is assumed to need in order to tackle certain test items, rather than on the actual processes which a reader might be expected to apply in real-world reading. The concern in this psychometrically driven approach is thus not with the actual components of the reading *process* that are necessary for comprehension, but with the factors which can be shown statistically to contribute to successful performance. The approach might be described as focusing upon a product in the form of the outcome of a test rather than upon the process which gave rise to it.

Many of these *post hoc* quantitative studies are limited to the extent they do not test the range of *types of reading* (careful and expeditious), nor do they consider the need to shape reading to the reader’s goals, or the *level of cognitive demand* imposed on the reader by a particular task.

Given the aim of evaluating the cognitive validity of reading tests, an approach premised solely on a *post hoc* factorial analysis of reading tests seems problematic. Weir (2005:18) cautions against relying on this procedure for construct validation as ‘statistical data do not in themselves generate conceptual labels’. Field (in preparation) echoes this position, pointing to the ‘dangers of relying exclusively on an approach that attempts to track back from a product or outcome to the process that gave rise to it.’ Such analyses by their nature tell us little about what is actually happening when a reader processes text under test conditions. We need to go deeper and examine as far as is possible the nature of the reading activities in which we engage during a test in such a way as to enable comparison with activities occurring during non-test reading. We argue below that in respect of cognitive validity we need to establish clearly the types of reading we wish to include.

Informed tuition: a subskills approach to defining reading comprehension

The kind of factorial approach described above emerged during a time when the climate of opinion in the methodology of teaching reading strongly favoured what

1. This article is based on material in preparation for *Examining Second Language Reading*, a forthcoming title in the Studies in Language Testing series; see www.cambridgeesol.org/what-we-do/research/silt.html for further information.

was termed a *subskills approach*. Like the factorial approach employed by researchers, it assumed that reading might be subdivided into the competencies which the skilled reader is believed to have. In L2 pedagogy, the development of the subskills movement (Grellet 1987, Munby 1978, Nuttall 1996) aimed to break reading down into constituent competencies. This arose in large part from the need to develop communicatively-oriented pedagogical syllabuses and the need felt by teachers to provide more focused practice in the skill as an alternative to relying on more and more general reading. The approach has mainly been based on informed intuition rather than empirical research but has been found to be useful by a generation of teachers. As a result it has figured prominently in EFL reading materials for teaching purposes and test specification (see Williams and Moran 1989). It became accepted pedagogical practice to break the reading process down and to address component skills separately.

In the field of testing, the subskills approach has given rise to the notion that it is possible to link particular item or task types to specific subskills that they are said to tap into. A growing body of literature (e.g. Bachman *et al.* 1988, Lumley 1993, Weir and Porter 1994) suggests that it is possible with clear specification of terms and appropriate methodology for testers to reach closer agreement on what skills are being tested. Similarly, Alderson (2005:125–137) in the DIALANG project noted that individual items are now viewed as testing identifiable skills.

However, the value of this subskills approach for testing is contentious. The jury is still out on whether it is possible for expert judges to be convincingly accurate in their predictions about what competencies individual items in a test are assessing. Test developers may be better served if they attempt to design the overall spread of items in a test in such a way as to cover the reading construct that is appropriate to reading purpose and target level of processing difficulty; if they identify which types of reading are most appropriate to different proficiency levels and attempt to ensure that the cognitive processing demands needed to complete such tasks are commensurate with the skilled reading process as evidenced by research in cognitive psychology.

Limitations of the informed intuition approach

Informed intuitive approaches have been helpful in advancing our conceptualisation of what is involved in reading both for pedagogical and assessment purposes. The problem is that they were more organisationally than theoretically driven; they often only represent the views of expert materials designers as to what is actually being tested in terms of reading types. More importantly, the central role of the *test taker* has been largely overlooked.

So far little reference has been made to the cognitive processing that might be necessary for second language (L2) candidates to achieve the various types of reading initiated by the reading test tasks employed. To clearly establish the trait that has been measured we need to investigate the processing necessary for task fulfilment which is the focus of the next section.

A cognitive processing approach to defining reading comprehension

In attempting to understand what is involved in the process of reading comprehension, researchers have proposed various theories and models of reading (e.g. Birch 2007, Cohen and Upton 2006, Goodman, 1967, Gough 1972, Just and Carpenter 1980, LaBerge and Samuels 1974; Kintsch and van Dijk 1978, Perfetti 1999, Rayner and Pollatsek 1989). These theorists all recognise the reading process as combining “bottom-up” visual information with the “top-down” world knowledge that the reader brings to the task; but they have diverged in their accounts of the importance accorded to each and of the ways in which the two sources of information are combined by the reader.

In bottom-up processing linguistic knowledge is employed to build smaller units into larger ones through several levels of processing: the orthographic, phonological, lexical, syntactic features of a text and then sentence meaning through to a representation of the whole text. In top-down processing larger units affect the way smaller units are perceived. Sources of information include context, where general and domain specific knowledge is used to enrich propositional meaning, and/or the developing *meaning representation of the text so far*, created in the act of reading a text.

There are two distinct uses for context: one to enrich propositional meaning extracted from a decoded text and the other to support decoding where it is inadequate. Stanovich (1980) argues that an interactive compensatory mechanism enables unskilled readers to resort to top-down processing through using contextual clues to compensate for slower lexical access due to inaccurate decoding. He suggests that skilled L1 readers employ context for enriching understanding rather than for supplementing partial or incomplete information as is the case for the poor reader. Jenkins *et al.* (2003) note research suggesting that skilled readers rarely depend on top-down prediction to identify words in context because they have such rapid word identification skills which outstrip the rather slower hypothesis forming top-down processes. The opposite is true for less skilled readers as their bottom-up processing of print is slower than top-down word prediction processes. The current accepted view, however, is that we process at different levels simultaneously and draw on both bottom-up and top-down processes in establishing meaning.

The cognitive validity of a reading task is a measure of how closely it elicits the cognitive processing involved in contexts beyond the test itself, i.e. in performing reading tasks in real life. We have drawn on the work of authors working within the field of cognitive psychology in order to devise a model of the L1 reading process – supported by empirical evidence – which can be treated as the goal towards which the L2 reader aspires.

There will of course be some individual variation in cognitive processing but we need to consider whether there are any generic processes that we would want to sample in our reading tests which would bring the process of comprehending in a test closer to that in real life. The generic cognitive processes contributing to reading that we

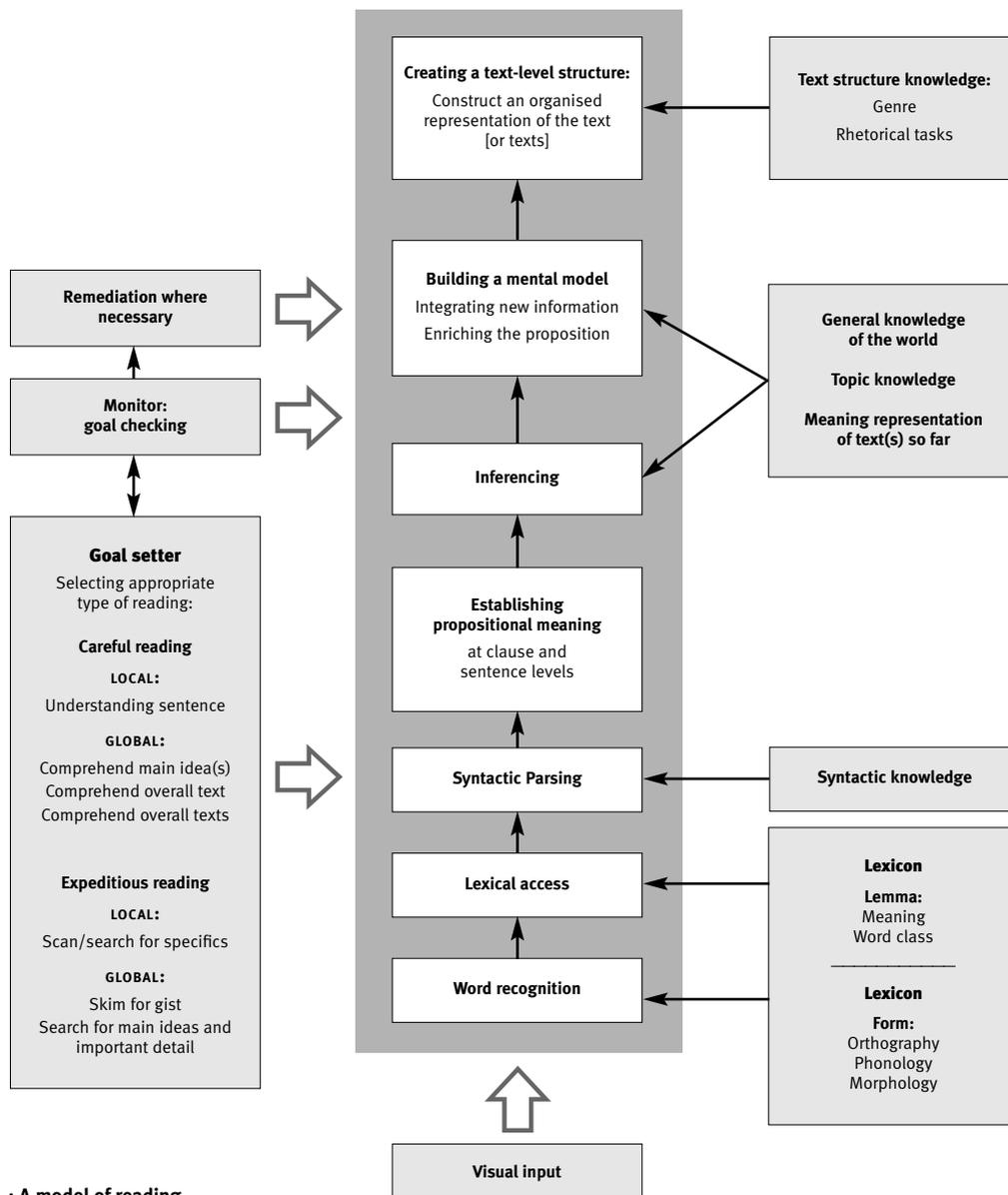


Figure 1 : A model of reading

have identified from the literature are represented in Figure 1 and explained in the subsequent text.

In discussing these cognitive processes, we will start with a brief description of the metacognitive activity of a *goal setter* (see left hand column) because, in deciding what type of reading to employ when faced with a text, critical decisions are taken on the level(s) of processing to be activated in the central core of our model. The various elements of this *processing core* (see middle column) which might be initiated by decisions taken in the goal setter are then described individually. A discussion of the *monitor* then follows as this can be applied to each of the levels of processing that is activated in response to the goal setter's instructions. We then return to discuss in more detail the *types of reading* we have listed under the *goal setter* and relate them to appropriate elements from the central processing core.

The Goal Setter

The goal setter is critical in that the decisions taken on the purpose for the reading activity will determine the relative

importance of some of the processes in the central core of the model. Urquhart and Weir (1998) provide an overview of the goals that are open to the reader and characterise reading as being either *careful* or *expeditious* or taking place at the *local* and *global* level.

Global comprehension refers to the understanding of propositions beyond the level of micro-structure, that is, any macro-propositions including main ideas, the links between those macro-propositions and the way in which the micro-propositions elaborate upon them. At the macro-structure level of the text the main concern is with the relationships between ideas represented in complexes of propositions which tend to be logical or rhetorical. Individual components of these complexes are often marked out by the writer through the use of paragraphs. This kind of process is important in *careful global reading* operations where the reader is trying to identify the main idea(s) by establishing the macro-structure of a text. It is also related to *search reading* where the reader is normally trying to identify macro-propositions but through short cuts due to time pressure. Global comprehension is also related

to the top structure level of the text where the reader, through *skimming* is trying to establish the macro-structure and the discourse topic and in *careful global reading* to determine how the ideas in the whole text relate to each other and to the author's purpose.

Local comprehension refers to the understanding of propositions at the level of micro-structure, i.e. the sentence and the clause. Cohen and Upton (2006:17) suggest that local comprehension is strongly associated with linguistic knowledge. Alderson (2000:87) makes the connection between local comprehension and test items which focus on understanding explicit information. In textually explicit questions, the information used in the question and the information required for the answer are usually in the same sentence. In our model above, local comprehension is at the levels of *decoding (word recognition, lexical access and syntactic parsing)* and *establishing propositional meaning* at the sentence and clause level.

Careful reading is intended to extract complete meanings from presented material. This can take place at a local or a global level, i.e. within or beyond the sentence right up to the level of the complete text. The approach to reading is based on slow, careful, linear, incremental reading for comprehension. It should be noted that models of reading have usually been developed with *careful reading* in mind and have little to tell us about how skilled readers can cope with other reading behaviours such as skimming for gist (Rayner and Pollatsek 1989: 477–478).

Expeditious reading involves quick, selective and efficient reading to access desired information in a text. It includes *skimming, scanning* and *search reading*. *Skimming* is generally defined as reading to obtain the gist, general impression and/or superordinate main idea of a text; accordingly it takes place at the global text level. *Scanning* involves reading selectively, at the local word level, to achieve very specific reading goals, e.g. looking for specific items in an index. *Search reading*, however, can take place at both the local and global level. Where the desired information can be located within a single sentence it would be classified as local and where information has to be put together across sentences it would be seen as global. In both cases the search is for words in the same semantic field as the target information unlike scanning where exact word matches are sought.

Once we have discussed the central processing core of the model from the bottom level upwards, we will return to these purposes for reading in order to examine the relationships between the intended purpose and the processing activity it elicits in this central core.

Central Processing Core

The processes described here attempt to characterise the reading behaviours available to the competent L1 reader which the L2 reader might be expected to progressively approximate to as their proficiency level in L2 improves. The knowledge base on the right hand side of the model is drawn upon by the central processing core in line with the intended purpose and the performance conditions established by the task.

Word recognition

Word recognition is concerned with matching the form of a word in a written text with a mental representation of the orthographic forms of the language. In the case of the less experienced L2 reader, the matching process is complicated by a more limited sight vocabulary in the target language, and by the fact that the reader does not make the kind of automatic connection between written word and mental representation that an experienced reader would. Field (2004:234) cites Coltheart's (1978) dual route model of decoding which suggests that we process written words in two ways. A *lexical route* enables us to match whole words while a *sub-lexical* route permits us to identify words by means of grapheme-phoneme correspondence. All languages appear to use both routes. The problem for the L2 reader of English is that it is much more difficult to match an unfamiliar written form to a known spoken one by means of the sub-lexical route or to internalise the spoken forms of written words. Much of the matching during the acquisition of L1 reading skills in English relies quite heavily on analogies between words with similar written forms (*light – fight – right*). L2 learners, with limited vocabulary and less automatic pattern recognition, are less able to apply these analogies.

The opaque orthography of English may result in greater dependence on the lexical route and thereby increase the difficulty when unskilled L2 readers meet words in text which they have never encountered before in a written form. This may mean that test developers need to ensure that at lower levels of proficiency the number of unknown words in a text need to be controlled and the length of texts these candidates are exposed to will need to be shorter than those for skilled readers. L2 readers with L1 language backgrounds in which the orthographies are very dissimilar to that of English, e.g. in the script or direction of reading, will face additional problems at the decoding level (see Birch 2007).

Jenkins et al. (2003) note that less skilled readers are constrained by inefficient word recognition which requires attentional resources and uses up available working memory capacity that might otherwise be used for comprehension. In the skilled reader, efficient word recognition frees up attentional resources thereby increasing the capacity in working memory available for more complex operations. Accuracy and automaticity of word recognition is critical for the skilled reader (see Grabe 2004, Perfetti 1997, Wagner and Stanovic 1996). Automaticity is the result of increasing experience in decoding and of the mind's orientation towards creating processes which are undemanding upon attention. Those readers who can decode accurately and automatically will backtrack less often and have more attentional capacity available in working memory for comprehension, e.g. establishing propositional meaning, inferencing and building a mental model and integration of information across sentences.

Lexical access

Field (2004:151) describes this as the 'retrieval of a lexical entry from the lexicon, containing stored information about

a word's form and its meaning'. The form includes orthographic and phonological mental representations of a lexical item and possibly information on its morphology. The lemma (the meaning-related part of the lexical entry) includes information on word class and the syntactic structures in which the item can appear and on the range of possible senses for the word. The orthographic form plays a part in what was described in the previous section as word recognition. Some accounts describe sets of visually similar words in the reader's mental vocabulary as being in competition with each other. Individual words are activated in relation to the extent to which they do or do not resemble a target word on the page. Finally, a point is reached where one word accumulates so much evidence that it is selected as the correct match.

Frequent words appear to be identified more quickly than infrequent ones because, according to serial models of lexical access, words are stored on this principle. Other theories such as parallel access suggest that words are *activated* in accordance with their frequency and the closest match to context (Field 2004:117,151). This suggests that test developers need to ensure that there is a suitable progression in terms of lexis from frequent words to those with less frequent coverage as one moves up the levels of proficiency in L2 reading examinations.

Syntactic parsing

Fluency in syntactic parsing is regarded as important in the comprehension process by a number of authorities (Perfetti 1997). Once the meaning of words is accessed, the reader has to group words into phrases, and into larger units at the clause and sentence level to understand the message of the text. Cromer (1970) illustrates the importance of competence in the syntax of the target language for deriving meaning from text. He demonstrates that good comprehenders use sentence structure as well as word identification to comprehend text (see also Haynes and Carr 1990). It is therefore important that test developers ensure that the syntactic categories appearing in texts employed at each proficiency level are appropriate to the candidate's level of development.

Establishing propositional (core) meaning at the clause or sentence level

Propositional meaning is a literal interpretation of what is on the page. The reader has to add external knowledge to it to turn it into a message that relates to the context in which it occurred.

Inferencing

Inferencing is necessary so the reader can go beyond explicitly stated ideas as the links between ideas in a passage are often left implicit (Oakhill and Garnham 1988:22). Inferencing in this sense is a creative process whereby the brain adds information which is not stated in a text in order to impose coherence. A text cannot include all the information that is necessary in order to make sense of it. Texts usually leave out knowledge that readers can be trusted to add for themselves. Problems may of course arise where the assumed knowledge relates to that of the L1 host culture and such inferences are not possible for the L2

learner who lacks this knowledge. Inferencing may also take place at word level, when a word is referring to an entity as in the case of pronouns or is ambiguous in its context or is a homograph. It may also involve guessing the meaning of unknown words in context.

Hughes (1993) argues that we should replicate real-life processes and attempt to sample all types of inferencing ability in our tests with the caveat of being able to select texts which are close to the background and experience of the candidature. However, he admits that pragmatic inferencing questions (where the reader not only makes use of information in the text but also refers to their own world knowledge) are problematic where candidates have very different knowledge, experience and opinions. Pragmatic evaluative inferences are particularly difficult to include in tests because of the marking problems associated with potential variability in answers. Even though the evidence available to readers from the text is given, they will come to it with different perspectives and expectations (see Chikalanga 1991 and 1992). Test developers need to be conscious of this at the item-writing stage to avoid penalising candidates who may lack particular world knowledge.

Building a mental model

Field (2004:241) notes that 'incoming information has to be related to what has gone before, so as to ensure that it contributes to the developing representation of the text in a way that is consistent, meaningful and relevant. This process entails an ability to identify main ideas, to relate them to previous ideas, distinguish between major and minor propositions and to impose a hierarchical structure on the information in the text.' Ongoing meaning representation is provisional and liable to revision as well as updating with new information from the text. *Selection* may occur whereby stored information is reduced to what is relative or important.

According to Kintsch and van Dijk (1978:374), the propositions representing the meaning of a text are linked together, usually by argument overlap, to form a hierarchical text base. Micro-structures are processed, converted into semantic propositions, and stored in the working memory, while the cohesion between them is established. As the process moves on, a macro-structure is built up. Background knowledge, stored in long term memory, is utilised to supply an appropriate schema for the macro-structure, as well as to aid coherence detection in the construction of the micro-structure. Crucial information tends to be at the top levels of this hierarchy, while detailed information is at the lower levels.

As we discuss below, while building a mental model there is a need to monitor comprehension to check the viability of the ongoing interpretation. Monitoring chiefly checks the consistency of incoming information against the meaning representation established so far. If the two conflict the reader regresses to check. This type of monitoring is especially absent in weaker readers. World knowledge in the form of schemata in long term memory plays an important part in judging the coherence and consistency of what has been understood when it is integrated into the ongoing meaning representation.

Creating a text-level structure

At a final stage of processing, a discourse-level structure is created for the text as a whole. The skilled reader is able to recognise the hierarchical structure of the whole text and determines which items of information are central to the meaning of the text. The skilled reader determines how the different parts of the text fit together and which parts of the text are important to the writer or to reader purpose.

The development of an accurate and reasonably complete text model of comprehension would therefore seem to involve understanding of discourse structure and the ability to identify macro level relationships between ideas. It also involves understanding which propositions are central to the goals of the text and which are of secondary importance.

The Monitor

Thus far we have looked at each of the levels of processing that may be brought into play as a result of metacognitive decisions taken in the goal setting stage. A further metacognitive activity may take place after activation of each level of the processing core: test takers are likely to check the effectiveness of their understanding (Sticht and James 1984). The *monitor* is the mechanism that provides the reader with feedback about the success of the particular reading process.

Self-monitoring is a complex operation which may occur at different stages of the reading process and may relate to different levels of analysis. In decoding text, monitoring involves checking word recognition, lexical access, and syntactic parsing. Within meaning building it can involve determining the success with which the reader can extract the writer's intentions or the argument structure of the text. Researchers like Perfetti (1999) or Oakhill and Garnham (1988) have argued that the unskilled L1 reader often fails to monitor comprehension or at least makes less use of monitoring strategies, particularly at the comprehension level. Studies have also shown that one of the hallmarks of a good reader is the ability to check the meaning representation for consistency. Skilled readers, on failing to understand a part of a text, will take action such as rereading to deal with the problem (see Hyoua and Nurminen 2006).

The components of goal setter and monitor can be viewed as *metacognitive mechanisms* that mediate among different processing skills and knowledge sources available to a reader. Urquhart and Weir (1998) provide detailed explanations about how these metacognitive mechanisms enable a reader to activate different levels of strategies and skills to cope with different reading purposes. The reader may choose to skim, search read, scan or read carefully in response to the perceived demands of the task. The level of processing required by the activity will also relate closely to the demands set by the test task.

Relating reading types to the central processing core

The goal setter part of the model is critical in that the decision taken on the purpose for the reading activity will determine the processing that is activated, dependent, of course, on the limitations imposed by the L2 reader's

linguistic and pragmatic knowledge and the extent of the reader's strategic competence. Rothkopf (1982) illustrates how the purpose for reading a text determines what and how much the reader takes away from it. Once the readers have a clear idea of what they will be reading for, they can choose the most appropriate process(es) for extracting the information they need in the text (see Pressley and Afflerbach 1995 for a comprehensive review of planning processes). The goal setter determines the overall goal of the reading, and also selects the type of reading which is likely to achieve that goal. Below we describe three expeditious reading types: skimming, search reading and scanning, and we discuss careful reading.

Expeditious reading: skimming

For Urquhart and Weir (1998) the defining characteristics of skimming are that the reading is selective and an attempt is made to build up a macro-structure (the gist) on the basis of as few details from the text as possible. Skimming is selective depending on how much information readers decide to process, they may access words and possibly process entire sentences. The reader will allocate his attention: focusing full attention on propositions that seem to be macro-propositional and reducing attention on others. He uses knowledge of text and genre which indicates likely positions for macro-propositions, e.g. first sentence of paragraph.

Presumably, the *monitor* checks as to whether the material surveyed is appropriate or not; in this case, the amount processed may be quite substantial. The reader will pause at appropriate points to semantically process words, phrases and clauses. If skimming is equivalent to gist extraction, then presumably propositions are committed to the long term memory on the hypothesis that they represent the macro-structure. That is, the process of debating whether a proposition is part of the macro-structure or not, which is assumed to take place during careful reading, is here replaced by a guess that it is usually supported by general knowledge of the world or domain knowledge.

The reader is trying to build up a macro-structure of the whole text (the gist) based on careful reading of as little of the text as possible. This is why skimming does not lend itself to the construction of numerous test items. A study of samples of EAP reading tests such as TEEP, or IELTS (see Weir et al. 2000) reveals that skimming rarely features in items in those tests and when it does, it is realised in only a single item asking a question such as 'What is the main idea of this passage?' (see TOEFL 1991 and UETESOL 1996).

Skimming requires the creation of a skeletal text level structure and, in particular, a decision as to the superordinate macro-proposition (Kintsch and van Dijk 1978) that encapsulates the meaning of a text. However, because of the rapid and selective nature of the processing involved it is unlikely to result in a detailed meaning representation of the whole text, a meaning representation that includes the relationships between all the macro-propositions and their relative importance. To arrive at a comprehensive and accurate text level structure, careful, rather than merely expeditious, global reading is necessary.

One can locate macro-propositions in two ways:

- by selective eye movements which attempt to locate sentences within the text stating the major issues
- by evaluating propositions as they are read, in order to identify those which start a new meaning structure rather than those that are subservient to or supportive of an ongoing structure.

Search reading

In search reading, the reader is sampling the text, which can be words, topic sentences or important paragraphs, to extract information on a predetermined topic. The reader may draw on formal knowledge of text structure to assist in this search for information on pre-specified macro-propositions (see Trabasso and Bouchard 2002, Urquhart and Weir 1998). Pugh (1978:53) states that in search reading:

'the reader is attempting to locate information on a topic when he is not certain of the precise form in which the information will appear... the reader is not pursuing a simple visual matching task (as in scanning), but rather needs to remain alert to various words in a similar semantic field to the topic in which he is interested. It is true that the visual activity involved is similar to scanning in many ways. However, the periods of close attention to the text tend to be more frequent and of longer duration and, since information is more deeply embedded in the text, there is more observance of the way in which the author structures his subject matter and, hence, the linearity and sequencing. Information about the structure of the text may be used to assist in the search.'

For Urquhart and Weir (1998) search reading involves locating information on predetermined topics so the reader does not have to establish an overall representation of the whole of the text as in skimming. The reader wants only the relevant information necessary to answer the set questions on a text. In cognitive terms it represents a shift from generalised attention to more focused attention.

The start of the process is to look for related vocabulary in the semantic field indicated by the task or item. Once the required information to answer a question has been quickly and selectively located, careful reading will take over and this may involve establishing propositional meaning at the sentence level, enriching propositions through inferencing, and it may require the reader to integrate information across sentences. In the test situation the wording of the questions does not usually allow the candidate simply to match question prompts to text and so lexical access is more demanding than in scanning tasks.

Search reading involves the different aspects of meaning construction up to and including the level of building a mental model, but it does not require the creation of a text level structure. The relative importance of the information in the text (micro- versus macro-proposition) is not an issue: all that matters is that the information has a bearing on the knowledge that is sought.

Scanning

Scanning involves reading selectively, to achieve very specific reading goals. It may involve looking for specific words or phrases, figures/percentages, names, dates, or specific items at the local word level. It is a perceptual recognition process which is form-based and relies on

accurate decoding of a word or string of words. Rosenshine (1980) defines it as involving recognition and matching. The main feature of scanning is that any part of the text which does not contain the pre-selected word, symbol or group of words is passed over. A low level of attention is accorded until a match or approximate match is made. The reader will not necessarily observe the author's sequencing by following the text in a linear way.

Here, very few components of our model are involved. Suppose at the lowest level, the goal has been set as scanning a text to find a reference to a particular author. In fact, it is arguable that only a limited amount of lexical access is required. Presumably little or no syntactic processing needs to be involved, no checking of coherence, and no attempt to build a macro-structure. There is usually no need to complete the reading of the sentence, or to integrate the word into the structure of preceding text. As a result, scanning involves none of the different aspects of meaning building that we have identified in our model.

Careful reading

We will focus here on the different aspects of meaning building upon which this reading type depends and will distinguish between processing that takes place at the local and at the global level. Careful local reading involves processing at the decoding level until the basic meaning of a proposition is established. Some local inferencing might be required to build a mental model at the enriched sentence level. However, it does not entail integrating each new piece of local information into a larger meaning representation. The defining features of careful global reading are that the reader attempts to handle the majority of information in the text; the reader accepts the writer's organisation and attempts to build up a macro-structure on the basis of the majority of the information received.

Careful global reading draws upon all the components of the model. The reader decides to read the text with a relatively high level of attention as, for example, for the study of a core textbook at undergraduate level. The goal setter sets this attention level not just for the reading operation but also for the monitoring that accompanies it. The reader would normally begin at the beginning of the text and continue through to the end, employing the processes detailed in the central core of the model above: integrating new information into a mental model and perhaps finally creating a discourse level structure for the text where appropriate to the reader's purpose.

A more demanding level of processing in careful reading would be required when establishing how ideas and details relate to each other in a whole text. The reader not only has to understand the macro- and micro-propositions but also how they are interconnected. This will require close and careful reading and perhaps even a rereading of the whole text or at least those parts of it relevant to the purpose in hand. Most likely all of the processing components listed in the central core of our model above will be required in this "reading to learn" activity where there is new as well as given information to be understood. Cohen and Upton (2006:17) describe this "reading to learn" process. With reference to the new TOEFL iBT, they state that:

'...according to the task specifications (ETS 2003) Reading to learn is seen as requiring additional abilities beyond those required for basic comprehension. Reading to learn questions assess specific abilities that contribute to learning including the ability to recognise the organisation and purpose of a text, to distinguish major from minor ideas and essential from nonessential information, to conceptualise and organise text information into a mental framework, and to understand rhetorical functions such as cause-effect relationships, compare-contrast relationships, arguments, and so on...'

In the real world, the reader sometimes has to combine and collate macro-propositional information from more than one text. The likelihood is that the process would be similar to that for a single text representation model, but that after reading one text, the knowledge (and perhaps linguistic) base will have been expanded as a result of the final meaning representation of the text being stored in long term memory.

The need to combine rhetorical and contextual information across texts would seem to place the greatest demands on processing (Enright et al. 2000:4–7). Perfetti (1997:346) argues that this purpose requires an integration of information in a text model with that in a situation model in what he terms a *documents model* which consists of: 'An Intertext Model that links texts in terms of their rhetorical relations to each other and a Situations Model that represents situations described in one or more texts with links to the texts'. This would require more demanding processing than the other reading activities described above including a greater level of global inferencing and text level structure building and perhaps necessitating regression across whole texts.

Leaving aside for the moment the cognitive load imposed by the complexity of the text employed in the test, one might argue that difficulty in processing is in large part a function of how many levels of processing in our model are required by a particular type of reading. This is an issue on which we do not have empirical evidence. However we might hypothesise that *mutatis mutandis* the following order of difficulty might well obtain in reading types. Starting with the easiest and ending with most difficult our best guess would be:

1. Scanning/searching for local information
2. Careful local reading
3. Skimming for gist
4. Careful global reading for comprehending main idea(s)
5. Search reading for global information
6. Careful global reading to comprehend a text
7. Careful global reading to comprehend texts

As with most scales, we can be reasonably confident of the positioning at the two extremes (2 is more difficult than 1 and 7 more difficult than 6 in the scale above). The middle three types of reading (3, 4 and 5) are a closer call and it is likely that attention to contextual parameters might be necessary to establish clear water between these levels. Ashton (2003) using six subtests of CAE Reading demonstrated that items 6 and 7 on the scale above were consistently more challenging than item 5. Rose (2006)

conducted research during the review of FCE and CAE into the amount of time needed for candidates to complete a careful global reading multiple choice (MC) item and an expeditious local multiple matching item. She found that a careful reading MC item needed more time to answer than an expeditious multiple matching item, thus indicating that it was worth more marks.

Conclusion

In this article we have presented and argued for a cognitive processing approach as the most tenable and productive theoretical basis for establishing what reading comprehension really involves. This is especially so given the discussed limitations of the factorial approach tradition and those of the reading subskills approach – an approach based largely on “expert judgement” that takes little account of the cognitive processes test takers actually employ.

We hope that this model and account of cognitive processing will provide a useful basis for establishing the cognitive validity of reading tests, i.e. the extent to which the tasks that test developers employ will elicit the cognitive processing involved in reading context beyond the test itself.

References and further reading

- Alderson, J C (2000) *Assessing Reading*, Cambridge: Cambridge University Press.
- (2005) *Diagnosing foreign language proficiency: the interface between learning and assessment*, London: Continuum.
- Ashton, M (1998) An investigation into the task types used in the reading paper of the Certificate of Advanced English Examination (CAE), unpublished MA thesis, University of Reading, UK.
- (2003) The change process at the paper level. Paper 1 Reading, in Weir, C and Milanovic, M, 121–174.
- Bachman, L F, Kunnan, A, Vanniarajan, S, and Lynch, B (1988) Task and ability analysis as a basis for examining content and construct comparability in two EFL proficiency test batteries, *Language Testing* 5, 128–159.
- Birch, M (2007) *English L2 Reading: Getting to the Bottom*, Mahwah, NJ: Lawrence Erlbaum Associates.
- Block, E (1986) The Comprehension Strategies of Second Language Readers, *TESOL Quarterly* 20 (3), 463–94.
- Chikalanga, I W (1991) *Inferencing in the reading process: a cross cultural study*, unpublished PhD thesis, University of Reading, UK.
- (1992) A suggested taxonomy for inferences for the reading teacher, *Reading in a Foreign Language* 8 (2), 697–709.
- Cohen, A D and Upton, T A (2006) *Strategies in responding to the new TOEFL reading tasks* (Monograph No. 33), Princeton, NJ: Educational Testing Services.
- Coltheart, M (1978) Lexical access in simple reading tasks, in Underwood, G (Ed.), *Strategies in information processing*, London: Academic Press, 151–216.
- Cromer, W (1970) The difference model: a new explanation for some reading difficulties, *Journal of Educational Psychology* 61, 471–483.
- Davis, F B (1968) Research in Comprehension in Reading, *Reading Research Quarterly* 3, 499–545.

- Enright, M, Grabe, W, Mosenthal, P, Mulcahy-Ernt, P and Schedl, M (2000) *A TOEFL 2000 framework for testing reading comprehension: a working paper*, Princeton, NJ: Educational Testing Services.
- ETS (2003) *Task Specifications for Next Generation TOEFL reading test*, unpublished manuscript.
- Field, J (2004) *Psycholinguistics: the Key Concepts*, London: Routledge.
- (in preparation) Cognitive Validity, in Taylor, L (Ed.) *Examining Speaking* (Studies in Language Testing), Cambridge: Cambridge University Press/Cambridge ESOL.
- Goodman, K S (1967) Reading: A Psycholinguistic Guessing Game, *Journal of the Reading Specialists*, 6 (4), 126–135.
- Gough, P B (1972) One Second of Reading, *Visible Language* 6 (4), 291–320.
- Grabe, W (2004) Research on teaching reading, *Annual Review of Applied Linguistics* 24, 44–69.
- Grellet, F (1987) *Developing Reading Skills*, Cambridge: Cambridge University Press.
- Guthrie, J T and Kirsch, I S (1987) Distinctions Between Reading Comprehension and Locating Information in Text, *Journal of Educational Psychology* 79, 220–227.
- Haynes, M and Carr, T H (1990) Writing system background and second language reading: a component skills analysis of English reading by native speakers of Chinese, in Carr, T H and Levy, B A (Eds) *Reading and Its Development: Component Skills Approaches*, San Diego, CA: Academic Press, 375–421.
- Hughes, A (1993) Testing the ability to infer when reading in a second or foreign language, *Journal of English and Foreign Languages*, 10/11, 13–20.
- Hyona, J and Nurminen, A M (2006) Do adult readers know how they read? Evidence from eye movement patterns and verbal reports, *British Journal of Psychology* 97, 31–50.
- Jenkins, J R, Fuchs, L S, van den Broeck, P, Espin, C and Deno, S L, (2003) Sources of individual difference in reading comprehension and reading fluency, *Journal of Educational Psychology* 95 (4), 719–729.
- Just, M A and Carpenter, P A (1987) *The Psychology of Reading and Language Comprehension*, Boston, Mass.: Allyn and Bacon.
- Kintsch, W and van Dijk, T A (1978) Toward a Model of Text Comprehension and Production, *Psychological Review* 85, 363–394.
- LaBerge, D and Samuels, S D (1974) Towards a Theory of Automatic Information Processing in Reading, *Cognitive Psychology* 6, 293–323.
- Lumley, T (1993) The notion of subskills in reading comprehension: an EAP example, *Language Testing* 10 (3), 211–234.
- Munby, J (1978) *Communicative Syllabus Design*, Cambridge: Cambridge University Press.
- Nuttall, C (1996) *Teaching Reading Skills in a Foreign Language*, London: Heinemann.
- Oakhill, J and Garnham, A (1988) *Becoming a Skilled Reader*, Oxford: Blackwell.
- Perfetti, C A (1997) Sentences, individual differences, and multiple texts: three issues in text comprehension, *Discourse Processes* 23, 337–355.
- (1999) Comprehending written language a blueprint for the reader, in Brown, C M and Hagoort, P (Eds) *The neurocognition of language*, Oxford: Oxford University Press, 167–208.
- Pressley, M and Afflerbach, P (1995) *Verbal Protocols of Reading: The Nature of Constructively Responsive Reading*, Hillsdale, NJ: Lawrence Erlbaum.
- Pugh, AK (1978) *Silent Reading*, London: Heinemann.
- Rayner, K and Pollatsek, A (1989) *The Psychology of Reading*, Englewood Cliffs, NJ: Prentice Hall.
- Rose, D (2006) *Report on Observations of FCE and CAE Time Trials*, Cambridge: Cambridge ESOL internal report.
- Rosenshine, B V (1980) Skill Hierarchies in Reading Comprehension, in Spiro, R J, Bruce, B C and Brewer, W F (Eds), *Theoretical issues in reading comprehension*, Hillsdale, NJ: Lawrence Erlbaum, 535–559.
- Rothkopf, E Z (1982) Adjunct aids and the control of mathemagenic activities during purposeful reading, in Otto, W and White, S (Eds) *Reading expository material*, New York: Academic Press.
- Schedl, M, Gordon, A, Carey, P A and Tang, K L (1996) *An Analysis of the Dimensionality of TOEFL Reading Comprehension Items*, (TOEFL Research Reports 53) Princeton, NJ: Educational Testing Services.
- Stanovich, K E (1980) Toward an Interactive Compensatory Model of Individual Differences in the Development of Reading Fluency, *Reading Research Quarterly* 16 (1), 32–71.
- Sticht, T G and James, J H (1984) Listening and reading, in Pearson, P D, Barr, R, Kamil, M L and Mosenthal, P B (Eds) *Handbook of Reading Research* Volume 1, New York: Longman, 293–317.
- Trabasso, T and Bouchard, E (2002) Teaching Readers How to Comprehend Text Strategically, in Block, C C and Pressley, M (Eds) *Comprehension Instruction: Research-Based Best Practices*, New York: The Guilford Press, 176–200.
- Urquhart, A H and Weir, C J (1998) *Reading in a Second Language: Process, Product and Practice*, London: Longman.
- Wagner, R and Stanovich, K (1996) Expertise in Reading, in Ericsson, K A (Ed.) *The road to excellence*, Mahwah, NJ: Lawrence Erlbaum, 189–225.
- Weir, C J (2005) *Language Testing and Validation: An Evidence-Based Approach*, Basingstoke: Palgrave Macmillan
- Weir, C J and Milanovic, M (2003), *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913–2002* (Studies in Language Testing Volume 15), Cambridge: Cambridge University Press/UCL.
- Weir, C J and Porter, D (1994) The Multi-Divisible or Unitary Nature of Reading: The language tester between Scylla and Charybdis, *Reading in a Foreign Language* 10, 1–19.
- Weir, C J, Yang, H, and Jin, Y (2000) An empirical investigation of the componentiality of L2 reading in English for academic purposes, Cambridge: Cambridge University Press.
- Williams, E and Moran, C (1989) Reading in a foreign language at intermediate and advanced levels with particular reference to English, *Language Teaching* 22, 217–228.

Applying a cognitive processing model to Main Suite Reading papers

CYRIL WEIR UNIVERSITY OF BEDFORDSHIRE
HANAN KHALIFA RESEARCH AND VALIDATION GROUP

Introduction

In this article we apply the model of reading types and associated cognitive processes described in Weir and Khalifa (2008) to our Main Suite General English examinations.¹ Because of limitations of space we are only able to make detailed reference to two exams at adjacent proficiency levels – PET at B1 level and FCE at B2 level of the CEFR – in order to exemplify:

- the variety and complexity of the *reading types* demanded at B1/B2 levels
- the comprehensiveness of the *cognitive processes* covered by these two levels
- the cognitive demands imposed by relative *text complexity* in PET and FCE
- whether the cognitive processes elicited by these two exams resemble those of a reader in a non-test context.

This retrospective analysis of two long-standing examinations, which are supported by considerable expertise and experience in pedagogy, is a good basis for attempting to ground our theory in practice.

Cognitive Processing: Cambridge ESOL practice

In relation to levels, Cambridge ESOL examinations are aligned with the Common European Framework of Reference for Languages (CEFR; Council of Europe 2001; see Khalifa and Weir in preparation chapter 7 for a full exploration of the procedures involved in this). The CEFR refers to six reading stages for L2 learners of English. A1 and A2 stages describe the ability to read very slowly basic or straightforward information in a known area, i.e. very simple sentences or very short predictable texts. B1 level describes the ability to comprehend texts that consist of familiar or high-frequency everyday language; the CEFR states that learners operating at this level can ‘understand routine information and articles, and the general meaning of non-routine information within a familiar area’ and that at this level scanning for specifics introduces a variety in reading purpose and style and speed of reading for the first time (Council of Europe 2001). At B2 level, readers start focusing more on integrating the content of texts (e.g. the main ideas, the writer’s attitude, etc). Higher levels (C1 and C2) characterise more mature, proficient readers who are able to process more abstract texts with structurally and semantically complex language.

In many ways the CEFR specifications are limited in their characterisation of reading ability at the different levels and we need to be more explicit for testing purposes (see Weir 2005). We need to look closely at the types of reading targeted and the cognitive processes they activate. We must also consider the cognitive load imposed by the contextual parameters of the reading texts these activities are performed on.

In this article though we are limited to looking at only two adjacent levels (CEFR B1 and B2) in detail, we will make reference to these parameters across the proficiency range in our summary sections to give the reader a flavour of the wider picture.

The variety of reading types at B1/B2 level

Table 1 shows the variety of reading types and associated processing levels demanded at B1 and B2 levels which are described in detail in Weir and Khalifa 2008 (see Khalifa and Weir in preparation for further discussion of reading types at levels A2, C1 and C2).

Table 1: The variety and complexity of reading types at B1/B2 levels

	PET B1	FCE B2
Careful Reading Local		
Understanding propositional meaning at clause and sentence level	*	■
Careful Reading Global		
Comprehend across sentences	*	*
Comprehend overall text	■	(*)
Comprehend overall texts	■	■
Expeditious Reading Local		
Scanning or search reading	*	*
Expeditious Reading Global		
Skim for gist	■	■
Search reading	*	(*)

* Indicates a clear coverage of this type of reading

(*) Indicates only a limited coverage of this type of reading (1 or 2 items per part)

■ Indicates non-coverage of this type of reading

In PET the candidate has to cope with a range of both expeditious and careful reading types at both the local and the global level (see Table 2 for a description).²

1. All five levels of the Main Suite are described in *Examining Reading*, the Studies in Language Testing volume which the material in this article is adapted from (Khalifa and Weir in preparation).

2. This analysis is based on the past paper from PET December 2005 and the sample updated FCE materials. Whilst space does not permit reproduction of these Reading papers here, these can be viewed online at www.cambridgeesol.org

Table 2: Type of reading, task format and test focus in the PET Reading paper

Part	Type of reading	Format	Test focus
Part 1	Careful reading local	Three-option multiple-choice on 5 very short discrete texts: signs and messages, postcards, notes, emails, labels etc.	Reading real-world notices and other short texts for the main message.
Part 2	Expeditious reading global	Matching: 5 items in the form of descriptions of people to match to 8 short adapted-authentic texts.	Reading multiple texts for specific information and detailed comprehension.
Part 3	Expeditious reading local	True/False: 10 items with an adapted-authentic long text.	Processing a factual text. Scanning for specific information while disregarding redundant material.
Part 4	Careful reading global	Four-option multiple choice: 5 items with an adapted-authentic long text.	Reading for detailed comprehension; understanding attitude, opinion and writer purpose. Reading for gist, inference and global meaning.
Part 5	Careful reading local	Four-option multiple-choice cloze: 10 items, with an adapted-authentic text drawn from a variety of sources. The text is of a factual or narrative nature.	Understanding of vocabulary and grammar in a short text, and understanding the lexicogrammatical patterns in the text.

Table 3: Type of reading, task format and test focus in the updated FCE Reading paper

Part	Type of reading	Format	Test focus
Part 1	Careful reading global: tests candidates' ability to identify detail, opinion, gist, attitude, tone, purpose, main idea, meaning from context or text organisation features (exemplification, comparison, reference) in a text.	A text followed by 8 four-option multiple-choice questions.	Candidates are expected to show understanding of specific information, text organisation features, tone, and text structure in all parts.
Part 2	Careful reading global: tests candidates' ability to follow text development.	A text from which seven sentences have been removed and placed in a jumbled order, together with an additional distractor sentence after the text.	
Part 3	Expeditious local and occasional global reading: tests candidates' ability to locate specific information in a text or a group of texts.	A text or several short texts preceded by 15 multiple-matching questions.	

Expeditious reading appears in Parts 2 and 3 with an emphasis on scanning which tends to be the easiest of the reading types but also some search reading is intended in Parts 2 and 3. Due to logistical constraints, lack of control on time per part may mean that some candidates use careful reading rather than expeditious reading when completing these items.

In FCE the tasks focus on careful reading at the global as against the local level we noted above for PET (see Table 3 for a breakdown).

Expeditious reading appears in Part 3, but at FCE the emphasis is on the more complex search reading (local) rather than simply scanning. What also makes this part more difficult is that items are not in the right order and so more text needs to be processed than in the similar task in PET where items fall in accordance with the occurrence of information in the passage. Lack of control of time again, however, may mean that some candidates are given the opportunity to use careful reading rather than expeditious reading in completing these items.

Summary of reading types across all Cambridge Main Suite levels

Although in the preceding text we have limited ourselves to two adjacent proficiency levels for illustrating the

application of the socio-cognitive framework, we now provide a brief overview of Main Suite levels from A2–C2.

Overall there is a general progression of careful reading tasks in the Main Suite with local items only at KET and careful global items appearing for the first time at PET along with local items. At FCE only careful global items are tested.

The cognitive demands made on the candidate (i.e. the attentional resources demanded by the tasks) need to vary between the reading tests set at the different levels according to level of ability. This means that a task requiring understanding of text level representation may be less suitable below C1 level in the CEFR (CAE) because of the more demanding processing required for its successful completion. Considerations of 'cognitive demands' might lead us to conclude that items that demand text-level representation should be given much greater prominence in the more advanced Main Suite exams.

In terms of our cognitive processing model, we would argue that the ability to engage in higher level processing activities comes with a commensurate increase in language proficiency in the L2 (see Ashton 1998). Given the limited time and space available for testing reading skills and strategies it might be prudent to ensure that a test at the C2 level is eliciting data on the ability to cope with higher level processes. In any case to actually complete such tasks lower level processing must already be well automatised.

It is not possible to produce an overall representation of the whole text if decoding, understanding at the propositional level and building a mental model are beyond the ability of the candidate with regard to that particular text.

Skilled readers are more likely to recognise changes of topic in a text and to enrich their comprehension by bringing in general knowledge of the world or topic knowledge and build meaning at global (text) level rather than just at local (sentence) level. It may be that from the C1 level upwards we should expect readers to be able to answer items which test understanding of how ideas in a text relate to each other as the CEFR expects them to do.

It is argued in the literature that unskilled L2 readers are unable to adjust their processing modes when confronted with different purposes for reading (Koda 2005). It thus seems appropriate that at KET level (A2) the reader only has to process information carefully. The high incidence of expeditious items at PET level (B1) might be of some concern at this early stage but with the complexity of text being relatively simple (see Table 5 below) it appears appropriate in terms of cognitive load for B1 level. Given the fact that no time limits are placed on individual tasks the likelihood is that they may be processed carefully in any case.

In the CEFR there is no mention of different modes being applicable at levels A1–B1. However, the CEFR does state that at the B2 level readers are expected to be able to ‘scan texts for relevant information, understand detailed information and grasp the main point of a text’ (Council of Europe 2001). The expectation of a candidate at the C1 level is that they ‘can read quickly enough to cope with an academic course’, i.e. they can cope with expeditious global as well as careful global reading demands, adapting their reading style to meet different reading purposes (COE *ibid.*). There are no tasks which appear to test expeditious global reading at CPE and only a few items at CAE and FCE.

Oakhill and Garnham (1988:6) view skilled readers as being able to adjust their types of reading in line with the text they have to process and with what they wish to get out of it, e.g. skimming a newspaper versus processing a refereed journal article in your field of study. They note that skilled readers will do this “efficiently”. Block (1986:465–6) notes that a number of studies suggest that skilled L2 readers are better able to monitor their comprehension and select appropriate strategies flexibly according to type of text and purpose. This suggests that requiring candidates to adjust modes from task to task is appropriate when testing more advanced L2 candidates.

We have examined the types of reading that are represented in two levels of the Main Suite examinations. We will now turn to the central core of our model and look more closely at the level(s) of processing that appear to be necessary to cope with items in the various parts of each examination.

Levels of processing in B1 and B2 levels

In Table 4 we summarise those parts of the central processing core that appear to be elicited by the tasks in PET and FCE Reading papers.

Table 4: The range of cognitive processes covered by PET and FCE

	PET B1	FCE B2
Word recognition	*	*
Lexical access	*	*
Parsing	*	*
Establishing propositional meaning	*	*
Inferencing	*	*
Building a mental model	*	*
Creating a text level structure	■	■
Creating an organised representation of several texts	■	■

* Indicates a clear coverage of this type of reading

■ Indicates non-coverage of this type of reading

PET

- **Part 1:** Candidates establish propositional meaning at the sentence level in order to understand the meaning of a range of short, discrete texts. Inferences are sometimes required.
- **Part 2:** Selective scrutiny of the text is involved for each item as the candidate looks for words in the same semantic field as those in the question. When a potential match to the item is located an ongoing meaning representation is required as successful matching is dependent on a configuration of requirements across sentences in a short paragraph. The information is for the most part explicitly stated, but because several pieces of information are being sought at one time to make a decision on the correct answer, considerable demands are made upon working memory, thus increasing task difficulty.
- **Part 3:** Implicit in the presentation of questions before the passage is an assumption by the test designers that expeditious reading will be encouraged. The intention is not for students to read carefully, serially and incrementally, i.e. drawing on most elements of our processing model, although as the task is not constrained by time limits there is no guarantee that this does not happen. You do not have to read everything to answer the questions and this guidance is conveyed through exam reports, teaching resources website, and teachers’ seminars. Having the questions follow the order of the text makes the task easier as it assists the candidate in locating the necessary information more easily within the text and the candidate does not have to process text already covered in previous questions. The downside of this is that it rather compromises cognitive validity (in the interests of making the task easier) in that in real-life the points a reader wanted to check would not necessarily follow text order. Scanning is involved for exact matches in a number of items but otherwise the reader has to look for equivalent words in the text. The task requires occasional inferences.
- **Part 4:** Building a mental model is usually required but some questions can be answered within a sentence. In the latter case, within-sentence processing is less demanding because there is no necessity to make connections to build a wider meaning representation as

happens in building a mental model. Items which focus on understanding attitude, opinion and writer purpose may require integrating information across longer stretches of the text, for example where candidates are required to choose one of four possible summaries of the text or where they should read through the whole text first before answering a question on the purpose of the text.

- **Part 5:** Focus on lexical access and syntactic parsing. Establishing propositional meaning at the clause and sentence level.

In terms of processing there is an increase from KET in the number of items requiring integration of information across sentences and inferences are sometimes necessary to answer questions.

FCE

- **Part 1:** Usually requires integration of new information sometimes across large sections of the text. Many of the answers require the reader to form inter-propositional connections.
- **Part 2:** Requires integration of new information. No need to create a text level structure because sentences rather than paragraphs are being inserted. In order to complete the task successfully, candidates need to use the clues provided by, for example, discourse markers, understand how examples are introduced and changes of direction signalled. This often needs to be combined with inferencing.
- **Part 3:** Mostly only requires understanding sentence level propositions to answer the questions once the information has been located. May involve inferencing in those items which test understanding of attitudes or opinions.

In terms of processing there is a substantial increase over PET in the proportion of items dependent on the successful integration of information between sentences and many of these require inferencing.

Summary of cognitive processing across all Main Suite levels

The cognitive psychology literature tends to focus on learners at two ends of a skilled and unskilled spectrum – but how does this help to define intervening levels in the CEFR? The literature would seem to suggest that in general there is a progression in ability to cope with lower to higher level processing in the central core of our model as L2 reading ability develops.

Examinations in the Main Suite would seem to follow the order of difficulty in cognitive processing that is suggested by our model and the literature. The attentional resources of a reader are finite and, in the early stages of L2 development (A2 level candidates), one might expect a large part of those resources to be diverted towards more low-level considerations concerning the linguistic code. No matter what the L1, decoding processes are reliant upon recognising not only letters but letter clusters and whole words. This means that decoding at the level of form is bound to be problematic for any L2 reader – only assisted by the extent to which there are cognates in L2. Whereas in

L1 word recognition is highly automatic for practised readers, new form/meaning relationships need to be set up gradually for L2 and only slowly become automatised.

The effort of decoding makes considerable cognitive demands on the less skilled L2 reader and as a result is likely to become the principal focus of attention for many up to the A2 level and the main focus for tests set at these levels. There is often a failure to employ comprehension processes (e.g. using contextual information to enrich comprehension or higher level meaning building) partly because of the demands of decoding and partly (in the case of the L2 reader) because of the unfamiliar situation of reading a text where there are gaps in understanding and words and phrases are perceptually unfamiliar (see Perfetti 1985).

Textually implicit questions require the reader to combine information across sentences in a text and such questions are generally more difficult than explicit items based on a single sentence given the additional processing that is required (see Davey and Lasasso 1984). Oakhill and Garnham (1988) suggest that the less skilled reader fails to make a range of inferences in comprehension, from local links between sentences, to the way(s) the ideas in the whole text are connected. Hosenfeld (1977) likewise shows that use of inferencing strategy can discriminate between good and poor readers (see also Chamot and El-Dinary 1999).

Inferencing makes an appearance at A2 and B1 level in a few items at the sentence level but it is only at FCE (B2) and above that it begins to be tested widely and across larger areas of text especially in Part 1 at FCE. From FCE onwards, certain question types require the candidate to report not on information contained in the text but upon what that information entails.

Until L2 learners have relatively automatic processes for dealing with word recognition, lexical access and syntactic parsing, meaning-making beyond dealing with sentence level propositions is restricted. This is usually well established by the B2 level, when there is more processing capacity available in working memory for making propositional inferences, building a mental model and integrating information.

The ability to cope with questions requiring the candidate to develop an overall text representation of argumentative texts only takes place on reaching the C1 (CAE) level. The highest level of processing – that required to construct a text model of several texts – comes into play at the C2 (CPE) level. This would seem to be a reasonable way to define proficiency at these higher levels in terms of the cognitive processing involved.

Text complexity in Main Suite Reading papers

So far we have said little about the performance conditions, the contextual parameters under which reading activities take place. The complexity of the text is a function of how these contextual parameters are realised within it. Both individually and in combination they are likely to impact on the cognitive demands imposed upon the reader. A text with high frequency lexis is likely to be easier to process than a text of the same length on the same topic with a

large number of low frequency lexical items. A shorter text is likely to be easier to process than a significantly longer text *mutatis mutandis*. The calibration of a number of the key parameters affecting cognitive load across two Main Suite levels is relevant here and the impact of cognitive demands, lexical development and structural resources are discussed below.

Cognitive demands at B1/B2 levels

At PET level lexis is familiar and structures mainly simple and easy to parse. Propositional load is quite low and inter-sentence relationships are quite simple. Meanwhile, at FCE level the cognitive load is increased by the use of a broader range of vocabulary, some of which may be unknown to the candidate or less familiar, sentence structure and propositional content is more complex and text length is greater. The range of patterns from simple to complex at FCE as against mostly simple sentences at PET, and total text lengths amounting to 2000 words as versus approximately 1500 at PET add to the increase in cognitive demands between these two adjacent levels in the Main Suite. See Table 5 for an explanation of the cognitive demands imposed by relative text complexity at the two levels under investigation.

We now turn to a consideration of lexical development in the Main Suite.

Lexical development in Main Suite Reading papers

There would seem to be four key points to note with regard to lexical development in the Main Suite examinations. Firstly, it is inevitable that as candidates progress up the levels of the Main Suite, the lexical demands that are put upon them are stronger. The amount of less frequent, less well known vocabulary increases. Additionally, the number and complexity of the items that candidates are required to understand increases by level. Secondly, lexis at lower levels is restricted to everyday, literal and factual language. As students advance in proficiency, they are gradually expected to deal with increasingly subtle uses of the language of feelings and ideas. The senses associated with the words are less concrete and issues of polysemy may arise. More abstract texts will not be presented to candidates until levels C1 (CAE) and above. Thirdly, fiction inevitably requires a broader receptive vocabulary and this is introduced from FCE onwards taking the vocabulary beyond the candidate's repertoire. Lastly, from FCE upwards the extent to which the text is beyond what the reader knows increases. By CPE the candidate may be exposed to texts on any subject.

Structural resources in Main Suite Reading papers

The key points which relate to the structural resources used in reading texts in Main Suite practice are:

- An analysis of the papers across all five levels shows a very clear progression in terms of sentence structure from short, simple sentences to long, complex sentences. This is mirrored in the length of the texts used as very short texts are used at KET level and increasingly longer ones are employed at higher levels.
- This structural progression does not mean that some short sentences may not pose considerable difficulty and so still have a place in higher level texts. Ellipsis and colloquial use of language may make for short sentences that are hard to process and so only appropriate at more advanced levels.
- An increasing complexity of verb forms is also noticeable in texts as we move up the Cambridge ESOL levels. The use of modals, conditionals, inversion and other structures become more common as the texts used in the examinations become more concerned with conveying feelings and opinions, persuading and hypothesising rather than dealing simply with information as they do at lower levels.
- As well as sentence length and verb form, referencing is an aspect of structure that becomes noticeably complex in higher level texts where a reader needs to engage in quite complex anaphoric resolution and be aware of the contribution of synonyms to text coherence.

In addition, as one progresses up the levels propositional density and the complexity of relationship between propositions increases and adds to the cognitive load.

Conclusion

In this article we have applied the model of reading types and associated cognitive processes described in Weir and Khalifa 2008 to two of the Cambridge ESOL Main Suite exams. We have shown that in general across the suite, the range of careful and expeditious reading types we established in our model are covered appropriately, although there are a few anomalies at higher levels that may merit consideration. The reading types can be roughly calibrated to reflect the demands they make upon the candidate in terms of the levels of language processing upon which they draw. The processing necessary for these reading activities can be seen as a cline from decoding through the various layers of meaning construction as we move upwards in proficiency.

In grading the specifications for the five levels of the

Table 5: The cognitive load imposed by relative text complexity at B1/B2 levels

	Overall Number of words	Time allowed	Lexis	Structure
PET	1460–1590 words	35 items with a recommended 50 minutes	Everyday vocabulary of native speakers using English today plus lexis appropriate to candidates' personal requirements, for example, nationalities, hobbies, likes and dislikes.	Mostly simple sentences but some use of relative and other subordinate clauses.
Updated FCE	Approximately 2000 words	30 items administered in 60 minutes	Good range of vocabulary. Topics are addressed in detail and with precision.	A range of sentence patterns – from the simple to the complex.

suite, careful thought has been given to the relative cognitive difficulty both of the tasks and of the texts employed. Text demands are increased only gradually; and the more demanding types of reading, for example reading to comprehend the whole text and integrate information across texts, are reserved for higher levels of the suite.

Our analysis indicates that the Cambridge ESOL Main Suite examinations correspond closely to what we know of the cognitive processes involved in reading in real-life. The cognitive requirements have been adequately graded in relation to the different levels of the suite. Due consideration has been given both to task demands and to the types of processing that can be deemed to be representative of performance at different stages of proficiency.

References and further reading

- Ashton, M (1998) An investigation into the task types used in the reading paper of the Certificate of Advanced English Examination (CAE), unpublished MA thesis, University of Reading, UK.
- Block, E (1986) The Comprehension Strategies of Second Language Readers, *TESOL Quarterly* 20/3, 463–494.
- Chamot, A U and El-Dinary, P B (1999) Children's learning strategies in language immersion classrooms, *Modern Language Journal* 83, 319–338.

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.

Davey, B and Lasasso, C (1984) The interaction of reader and task factors in the assessment of reading comprehension, *Experimental Education* 52, 199–206.

Hosenfeld, C (1977) A Preliminary Investigation of the Reading Strategies of Successful and Nonsuccessful Second Language Learners, *System* 5 (2), 110–123.

Khalifa, H and Weir, C J (in preparation) *Examining Second Language Reading*, Studies in Language Testing, Cambridge: Cambridge University Press/Cambridge ESOL.

Koda, K (2005) *Insights into second language reading*, New York: Cambridge University Press.

Oakhill, J and Garnham, A (1988) *Becoming a Skilled Reader*, Oxford: Blackwell.

Perfetti, C A (1985) *Reading Ability*, New York: Oxford University Press.

Weir, C J (2005) Limitations of the Council of Europe's Framework of reference (CEFR) in developing comparable examinations and tests, *Language Testing* 22 (3), 281–300.

Weir, C J and Khalifa, H (2008) A cognitive processing approach towards defining reading comprehension, *Research Notes* 31, 2–10.

A corpus-informed study of specificity in Financial English: the case of ICFE Reading

ANGELA WRIGHT ASSESSMENT AND OPERATIONS GROUP

Introduction

This article reports on research into the characteristics of the core language of Financial English, and how it compares to Business English and General English. This study was completed in partial fulfilment of a Masters degree in Applied Linguistics at Anglia Ruskin University (UK). In Ingham and Thighe's report on the initial trialling of the International Certificate in Financial English (ICFE) in 2006, they state that 'background or subject knowledge, as expressed through work experience, may contribute to an increase in performance in ICFE in the [...] Test of Reading' (Ingham and Thighe 2006:8). It is, therefore, important to have a greater awareness of the characteristics of Financial English that make it specific.

Nation and Kyongho (1995) carried out a study that attempted to identify the point at which learners should start learning special purposes vocabulary, suggesting that a 2000-word general purpose vocabulary should be learnt first, but for those with special interests, it is best to then specialise in order to gain maximum benefit from text coverage. The question for test developers, as Read puts it, is 'how to identify ... lexical items for a context-specific vocabulary list' (Read 2000:229). Nation (2001) suggests

that this should be done by looking at the frequency and range of words, and identifying words of high frequency and with a wide range within the relevant area of specialisation. Clearly the simplest way to do this is by using corpus-informed techniques.

Methodology

The research outlined in this article was carried out using *WordSmith Tools* version 3.0 (Scott 1999) software. Three small corpora were created of 20,000–30,000 words each:

- a Financial English corpus based on ICFE Reading papers (covering Common European Framework – CEFR – levels B2 and C1)
- a Business English corpus consisting of reading texts from BEC Vantage (CEFR level B2) and BEC Higher (CEFR level C1)
- a General English corpus consisting of reading texts from FCE (level B2) and CAE (level C1).

WordSmith was used in three ways. Firstly, the *WordList* tool was used to create frequency lists for each corpus, which show all the words in the corpus in order of frequency and

the number of times they appear, and to run statistical analyses of features such as word and sentence length. Secondly, the *KeyWords* tool was used to identify words which are unusually frequent or infrequent, i.e. found more or less frequently in one corpus than another. Lastly, the *Concord* tool, which finds all the examples of a word in the corpus and shows them in concordance lines, was used to see how certain words are used in context, to identify common collocations, and to help ascertain whether words are used in a general or sub-technical sense. Various corpus-based studies in language for specific purposes, such as Flowerdew (2001), have found that sub-technical words, that is words that are in general usage but that have a special meaning when used in a technical sense (e.g. *balance, interest, margin*), are much more frequent than technical ones.

Reaching agreement on how to categorise words, however, is not straightforward. Horner and Strutt (2004) carried out research using data from the Cambridge Learner Corpus (CLC) from candidates' written responses to BEC exams. Raters were asked to categorise a list of items from the CLC into three groups according to their level of specificity. There was little agreement among the raters, leading Horner and Strutt to conclude that there is considerable 'fuzziness' involved in trying to categorise domain-specific lexis. Only seven words were considered by the majority of raters to be highly specific to a business context, and interestingly, six of these are related to finance: *audit, auditor, bookkeeping, accountant, banking* and *capitalisation*. The seventh word was *CEO*.

Nelson (2000) used *WordSmith Tools* to examine how Business English lexis differs from General English. He found that words that could be considered business-related represented only 1.41% of his total Business English corpus, but argues that the key words (words which appear with unusual frequency) found in his study represent lexis that is core to Business English. He concludes that Business English differs from General English, not only lexically, but also semantically, and to a certain extent grammatically.

It should be noted that for this study the lists were not lemmatised. Cover sheets and instructions to candidates were removed from the tests, but the task rubrics were not. Some words which appear regularly in rubrics, such as *example, question, page*, etc., were, therefore, disregarded when analysing the data. The findings from the three WordSmith analyses are presented and discussed below.

Word list analysis

WordList was used to compare features such as word and sentence length (see Table 1).

In the sample analysed, average word and sentence length in Financial English is longer than in both Business and General English, with there being more words of 10 letters or more than in Business English, which in turn has more long words than General English.

WordList was also used to produce lists of the most frequent words found in each corpus (see Table 2).

The two most frequent content words in both the Financial English and Business English corpora are *company/companies* and *business*, and there is

Table 1: Statistical analysis of the three corpora

	Financial English	Business English	General English
Tokens	20189	24010	30840
Average word length	5.12	4.67	4.45
Sentence length	26.04	22.72	24.93
9-letter words	1220	1241	1118
10-letter words	887	731	691
11-letter words	587	435	379
12-letter words	267	215	196
13-letter words	189	167	115
14-letter words	52	39	21

Table 2: Top 20 most frequent words in the three word lists

Financial word list		Business word list		General word list	
Word	Freq	Word	Freq	Word	Freq
business	102	company	157	people	96
financial	97	business	98	time	78
management	86	companies	80	work	57
company	84	people	75	house	52
information	59	management	64	book	46
risk	56	new	59	new	43
accountants	54	staff	47	design	41
market	49	year	46	well	41
companies	48	managers	43	life	39
accounting	47	customers	38	world	38
cost	42	time	38	good	36
value	40	sales	35	years	36
managers	39	years	33	reefs	34
capital	38	marketing	31	fish	31
share	35	products	30	day	30
investors	34	market	28	know	30
audit	33	brand	26	things	30
markets	33	knowledge	26	company	29
risks	32	manager	26	maze	29
time	32	process	26	see	28

considerable overlap in the Financial and Business frequency lists, but there are some exceptions. The words *financial, information, risk(s), accountant(s), accounting, cost, share, investor(s), audit* and *market(s)* all appear much further down the Business frequency list than the Financial one; whereas *staff, customer(s), sales, marketing, product(s)* and *brand* are much further up the Business list than the Financial one. Reasons for some of these differences are expanded upon further below, but there is a clear financial meaning to the majority of the more frequent words on the Financial list and a non-financial business meaning to most of the more frequent words on the Business list. There is a greater overlap between the Financial and Business lists, with seven words appearing in the top 20 words on both lists (*business, management, company, market, companies, managers, time*), than between Financial and General (two words: *company, time*). Five words appear in both the Business and General top 20 words (*company, people, new, years, time*). All the top 20 most frequent words in both the Financial and Business corpora were found to occur across a range of texts.

However, in the General corpus a few words, such as *design*, *reefs*, *fish* and *maze*, are limited to just one or two texts and so cannot be considered representative.

Key words analysis

Key words are words which appear in a text with greater or lower frequency than chance alone would suggest. Those occurring with greater than normal frequency are known as positive key words, and those with lower than normal frequency are negative key words. They are identified by comparing two corpora.

A *KeyWord* comparison of the Financial and Business corpora produced only 17 positive key words and two negative key words (Table 3), suggesting again that there is a strong overlap between the Financial and Business corpora.

Table 3: Key words in the Financial and Business word lists

Word	Freq	Financial list %	Freq	Business list %	Keyness
Positive key words					
financial	97	0.48	15	0.06	82.6
accountants	54	0.27	1		76.1
and	595	2.95	424	1.76	68.6
accounting	47	0.23	1		65.4
risk	56	0.28	8	0.03	49.5
audit	33	0.16	1		44.1
information	59	0.29	12	0.05	42.8
risks	32	0.16	1		42.5
reporting	31	0.15	1		41.0
our	98	0.49	39	0.16	37.9
budgeting	24	0.12	0		37.7
currency	23	0.11	0		36.1
environmental	20	0.10	0		31.4
MBA	22	0.11	1		27.6
cost	42	0.21	10	0.04	27.3
SSC	16	0.08	0		25.1
budgets	20	0.10	1		24.6
Negative key words					
his	7	0.03	66	0.27	45.3
he	13	0.06	92	0.38	53.9

As four of the positive key words (*currency*, *environmental*, *MBA*, *SSC*) appear frequently in just one or two texts each, they cannot be considered representative, as further indicated by their low keyness scores. The four functional words that appear on the list (two positive: *and*, *our*; and two negative: *his*, *he*) merit further investigation using *Concord* to see exactly how they are used. Of the remaining eleven positive key words, seven can be said to be clearly related to the world of finance and these may be the first words that can be identified as core financial lexis; they are: *financial*, *accountant(s)*, *accounting*, *audit*, *budgeting*, *cost* and *budget(s)*. The other positive key words (*risk(s)*, *information*, *reporting*) do not have such an obvious financial meaning and require further investigation using *Concord* to see how they are used in context.

Comparing the Financial corpus to the General corpus produced a much longer list, with 66 positive key words

and 19 negative key words. The positive key words list includes the same seven words that were identified as being core in the Financial/Business key word list (*financial*, *accountant(s)*, *accounting*, *audit*, *budgeting*, *cost(s)*, *budget(s)*), as well as the three others that require further investigation (*risk(s)*, *information*, *reporting*), and the two functional words (*and*, *our*), with *his* and *he* appearing once again on the list of negative key words. Of the remaining words, several seem to belong clearly to the world of finance and can also be said to form part of the core financial lexis: *investors*, *capital*, *value*, *finance*, *share*, *growth*, *profit*, *cash*, *investment*, *exchange*, *accountancy*, *accounts*, *price*, *fiscal* and *assets*. However, some of these are possibly sub-technical words, such as *capital*, *value*, *share*, *growth* and *exchange*, and *Concord* may be able to shed light on whether they are used with different meanings or as different parts of speech in Financial and General English, or with different collocations.

The positive key words list also includes six of the seven words which appeared in both the top 20 most frequent Financial and Business word lists (*business*, *management*, *company*, *market*, *companies*, *manager*) and many of the remaining words appear to be just as likely to be found in a business context as a financial one, including: *organisation(s)*, *service*, *corporate*, *target(s)*, *firm(s)*, *manufacturing*, *objective(s)* and *percent*. Indeed, they are all found in the Business corpus at least six times, although the collocation *corporate governance* only appears in the Financial corpus. *Concord* may also give insights into why some of the other words, which do not seem to have any obvious connection to finance, such as *control*, *statements*, *internal*, *efficiency* and *based*, appear on the list of positive key words. Most of the words on the negative key words list are functional rather than content words and require further investigation using *Concord*. Four categories stand out: personal pronouns, conjunctions; contractions and auxiliaries.

Using *KeyWord* to compare the Business and General corpora produced 40 positive key words and 9 negative ones. With the exception of *its* and *will*, all the words on the positive key words list are clearly related to business, although a few, such as *organisations*, *service* and *association*, may be used in non-business contexts. The inclusion of words such as *marketing*, *brand* and *HR* on this list suggests that they may come from areas of business that have little overlap with the financial world, and each of these words appears just once or not at all on the Financial list, although they do not appear as key words in the Financial/Business key word comparison.

Concordance analysis

Having identified the most frequent words using *WordList* and the key words using *KeyWord*, *Concord* was used to see how some of the most frequent and most 'key' words are used in context. Collocations of the words which appear in the top 20 most frequent words in both the Financial and Business lists were compared but no real differences in usage were found.

Concord was then used to see if various key words in the Financial/Business key word list and the Financial/General

key word list are used differently across the different corpora. As outlined above, there are three words that appear as positive key words on both lists, but which are not obviously connected to finance (*risk(s)*, *information*, *reporting*). While little discernable difference was found between how the word *information* was used across the different corpora and there were not enough examples of *reporting* in the Business or General corpora to compare usage, the word *risk(s)* appeared to be used slightly differently in Financial and General English. There are only five examples of *risk* in the General English corpus, which is too few to come to any conclusions, but four of these examples involve taking risks with the suggestion that this is positive or neutral, for example:

- *still young enough to take the risk*
- *risk-taking happens every day*

None of the 88 examples of *risk* found in the Financial texts, however, collocates with *take*, while many refer to *managing*, *avoiding* or *reducing* risk, with risk being considered something negative:

- *investments are considered high-risk*
- *taking action to lessen the risk*

Risk management, which appears 11 times, together with two examples of *managing risk*, is a particularly common collocation:

- *efforts to encourage sound risk management practices*

The *KeyWord* search also helped to identify a group of words which may be sub-technical and used differently in Financial and General English (*capital*, *value*, *share*, *growth*, *exchange*), so *Concord* was used to see if any differences could be found. All the examples of *capital* in the Financial corpus used capital in the financial sense. Of the two examples in General English, one referred to *venture capital* and the other to a *capital city*, suggesting that *capital* can be considered a sub-technical word and part of the core Financial lexis. *Value* was found 40 times in the Financial corpus and was always used in a financial sense. Of the five examples of *value* in the General English corpus, only one had a financial meaning. *Concord* also revealed that *share* is used differently in Financial from General English with 32 of the 35 examples referring to shares in a company, the majority of them to *share price*. Only one of the six examples found in the General English corpus relates to a '*share in the company*' and four of the remaining examples use *share* as a verb. There are not enough examples of *growth*, *exchange* or *statements* in the General corpus to compare usage, but all examples found of these words in the Financial corpus are clearly related to finance, so these may also form part of the core lexis, particularly the collocations *financial statements*, *exchange rate*, *stock exchange* and *foreign exchange*.

For the most part, there are no clear reasons why *control*, *statements*, *issues*, *internal*, *efficiency*, *key*, *systems*, *annual*, *based*, *basis* and *framework* appear as positive key words in the Financial corpus and there are not enough examples in the General corpus to make a comparison. A few collocations, however, stand out, including: *internal*

audit, *internal control*, *market efficiency*, *annual report* and expressions such as *knowledge-based economies*, *zero-based budgeting*, etc.

We now return to the analysis of key words of four particular types: personal pronouns, conjunctions, contractions and auxiliaries.

Personal pronouns

Our featured on both the Financial/Business and Financial/General key word lists as a positive keyword. *Concord* showed that *our* was found to refer exclusively to companies in both the Financial and Business corpora. It appears in a range of different texts in the Financial corpus, but seems to be strongly affected by text type in that 35 of the 98 instances of *our* come from just one text 'Letter to shareholders' and a further 15 come from 'Audit letter'. The other text types where it featured prominently were advertisements and company annual reports. In the Business corpus, a large number of the examples of *our* also appear in annual reports and similar financial reporting documents. However, *our* is used very differently in the General corpus. While there are six examples of *our* used to refer to a company, five from the same text, *our* is used to refer to a variety of groups of people (a family, a band, a group of friends, a group of students, etc.). However, it is also used in 11 instances to refer to people in general, for example:

- *In our 20s, we tend to socialise most nights*
- *The most glorious natural phenomena on our planet*
- *We ignore them at our peril*

We does not appear on either key word list, and is, in fact, more frequent in the General list (122) than Financial (93) or Business (50). Its usage, and also that of *us*, however, is very similar to *our*.

Both *he* and *his* appear as negative key words in Financial compared to Business and General English. Both words occur with a reasonable spread of frequency in the Business English corpus, but particularly in texts describing the careers of businessmen. The samples found in the General English corpus show that *he* and *his* are most frequently used in narratives and biographies of famous people. The use of both words is very limited in the Financial corpus and appears to be used mostly to quote the opinions of experts. What is most noticeable about the Financial and Business corpora compared to General English, however, is the difference in gender. *She* appears only three times in the Financial English corpus and five times in Business English, compared to 13 and 92 examples of *he*. *Her* does not appear on either list. In General English, *she* is still found less frequently than *he* (122 times compared to 171, with *her* occurring 110 times, *his* 93 times and *him* 43 times), but the gender difference is much less marked. Whether this is a reflection of the worlds of business and finance, a bias in the selection of texts, or both, cannot be established by this research, but Nelson (2000) found similar results in his much larger Business English corpus study, suggesting that the business world is more male dominated than the general one.

I and *my* appear as negative key words on the Financial/General list. *I* only appears 12 times in the

Financial list, and *my* and *me* three times each. The three pronouns are also fairly rare in Business (*I* 33 times, *me* 6 times, *my* 12 times) and they are used in quotes and in texts describing the careers of business people. *I* appears 265 times in the General list, *me* 33 times and *my* 80 times, mostly in narratives and life stories.

Although *it* appears as a negative key word on the Financial/General list, *Concord* did not help to identify any obvious reasons for this, suggesting that it is mostly used as an impersonal form (*it was still snowing, it is important that, it was not until*, etc.) in both corpora. On the other hand, *its*, which appears as a key word on the Business/General list and which appears 80 times in the Business list, 49 times in the Financial one and 35 times in the General list, is used almost exclusively to refer to companies in Business and Finance, but for a range of objects, books, animals, buildings, etc., as well as occasionally to companies, in the General one.

Conjunctions

The word *and* appeared on both the Financial/Business and Financial/General key word lists but, exploring this further using *Concord* did not reveal any particular differences in usage across the three corpora. Although there were examples of common collocations, such as *profit and loss account* (4), *public and private sectors* (3), *debtors and creditors* (3), this did not stand out as a particular feature of finance. The only noticeable difference was that *and* was used much more frequently to begin sentences in the General English corpus than either Financial or Business, mainly in narratives.

But is listed as a negative key word. However, it appears to be used in the same way in both the Financial corpus and the General corpus.

Contractions

The appearance of *s* and *t* as negative key words in the Financial/General list led to a check on whether these were being used as contractions, possibly suggesting a greater degree of formality in Financial English. *T* was more revealing than *s*, as the latter was frequently used in genitive forms. Contractions using *t*, such as *don't*, *didn't*, *aren't*, appeared 102 times in General English, 15 times in Business English and only 8 times in Financial English. While the General corpus is 50% larger than the Financial one, this nevertheless suggests that there is more use of contractions in General English than Financial English. Once again, however, this may relate to the use of narrative texts in the General English tests.

Auxiliaries

Will appears as a positive key word in the Financial/General list, appearing 102 times in the Financial corpus compared to 51 times in the General one. In addition, there are nine examples of the contracted form *ll* in the General English list but none in Financial English. *Will* is fairly evenly distributed across a range of texts in the Financial corpus, but it shows certain patterns of usage. Firstly it appears quite frequently in letters and is used to explain procedures:

- *We will notify you immediately*
- *Our fees will be billed as work progresses*

It is also found in company annual reports to express intentions:

- *We will grow these businesses organically*
- *In the future, we will concentrate on ...*

Lastly, it is found in a range of articles with two different functions, predicting and describing general truths:

- *Difficulties with foreign exchange will probably have an influence*
- *Auditors will have to expect to report on ...*
- *Demand for the share will rise, and so its price will rise*
- *The strategic plan will normally include financial projections*

In General English, on the other hand, *will* appears to be used for a wide range of functions and no clear pattern of usage emerges. In Business English, the use of *will* was found to be similar to Financial English but there were no examples of *will* in letters. On the other hand, there were several examples of *will* in job advertisements to describe the qualities required of an applicant and the duties they will carry out:

- *The successful applicant will have...*
- *You will join a small team*
- *You will be involved in all aspects of ...*

The past auxiliaries *was* and *had* both appear much more frequently in General English than Financial English (*was* 262 times compared to 26, and *had* 116 times compared to 9). The examples shown by *Concord* reflect once again the use of narrative texts in CAE and FCE. In the Financial English corpus, however, there are no narrative texts and the few instances of *was* and *had* appear mainly in company annual reports or company histories:

- *Net income in fiscal 2003 was nearly 50 percent higher*
- *AEI Ltd was listed on the stock exchange*

Again Business English follows a similar pattern to Financial English, although *was* (74 examples) and *had* (38) are more common. The frequency of texts describing the life story of a business person may account for the difference.

Establishing the specificity of Financial English reading texts

Although the small sample of texts used could not establish 'core language' with any degree of certainty or to any depth, some noticeable differences were found between Financial and General English, and a strong overlap was found between Financial and Business English.

As many of the words on the Financial/General key word list are also found on the Business/General list, and there is considerable overlap in the 20 most frequent content words in both the Financial and Business corpora, this suggests that there is considerable fuzziness between Financial and Business English. A detailed comparison of how words were used in the Business and Financial texts

using *Concord*, found very little difference between the two domains. Although the study suggests that there may be areas of Business that have little or no overlap with Finance, such as marketing or HR, it would be difficult to state that there are areas of Finance that have no overlap with Business. Finance may, therefore, be a subset of Business, with certain core words that are found more frequently in this subset. Nevertheless, a number of words and collocations stand out as seeming to belong clearly to the Financial English domain, listed in Table 4.

Table 4: Proposed Financial English domain-specific vocabulary

Words	Collocations
accountancy, accountant(s), accounting, accounts	annual report
audit	corporate governance
assets	exchange rate
budget(s), budgeting	financial statements
capital	fiscal [+year]
cash	foreign exchange
cost(s)	internal audit
finance, financial	internal control
growth	risk management
investment, investor(s)	share price
profit	stock exchange
value	venture capital

While these words may form part of the core Financial lexis, this list is far from exhaustive. It seems very likely that many words that appear twelve times or fewer in the Financial frequency list, and, therefore, too infrequently to appear to be key, form part of the core lexis of Financial English, either as technical or sub-technical words, but the corpus used in this study was too small to establish this. These words include:

- fraud (12), invoice (12), quarter (12), shareholders/+ shareholder (12/6)
- profitability (11), balance (10), transactions/+ transaction (10/8), securities (9)
- liabilities (8), overheads (8), revenues (8), equity (7) etc.

Despite being part of the core Financial lexis, this does not mean that the words on this list are used exclusively in the Financial domain.

A further group of words was identified which belong to an overlapping Financial/Business core lexis, but which cannot be said to have any particular financial meaning, including:

- business(es), company/ies, corporate, firm(s)
- management, manager(s), manufacturing, market(s)
- objective(s), organisation(s), percent
- process(es), service, target(s).

Conclusion

In addition to the lexical differences between Financial, Business and General English described above, several other linguistic features were identified as being different,

including the length of words, use of contractions and degree of formality, and the use of conjunctions, personal pronouns and auxiliaries. Several of these differences were found to relate to the types of text used in the tests from the three different domains, particularly narratives and biographies. The inclusion of narratives in the General English tests helped to account for the increased use of the first person singular, contractions, past tense forms and possibly certain conjunctions. While genre analysis was beyond the scope of this research, it seems that some of the differences between Financial and General English may relate to genre, particularly the use of *our* in documents such as letters and annual reports, where a company may be describing its activities and/or achievements to shareholders, and the use of *will* for a range of functions in letters, annual reports and job advertisements.

The small size of the corpora used in this study raised a number of issues; it restricted the possibilities for making comparisons of how words were used differently in different corpora, and it led to some words appearing to be much more frequent than they would in a larger corpus. It also meant that it was only possible to identify the most common core financial terms, the majority of which a candidate at CEFR level B2/C1 taking a General English test would probably be very familiar with, even if they are likely to come across them less frequently than a finance/accounting student or professional. Hargreaves (2000) stated that Cambridge ESOL's approach to testing lexis was corpus-informed but not corpus-based. This will no doubt remain the case, but for LSP exams such as ICFE, further larger corpus studies are necessary to build a clearer picture of the nature of the target domain. This study is a small but important step towards achieving that goal.

References and further reading

- Flowerdew, J (2001) Concordancing as a tool in course design, in Ghadessy, M, Henry, A, Roseberry, R (Eds), 71–92.
- Ghadessy, M, Henry, A, Roseberry, R (Eds) *Small Corpus Studies and ELT: theory and practice*, Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hargreaves, P (2000) Collocations and testing, in Lewis M (Ed.), *Teaching Collocation: Further Developments in the Lexical Approach*, Hove, UK: Language Teaching Publications, 205–223.
- Horner, D and Strutt, P (2004) Analysing domain-specific lexical categories: evidence from the BEC written corpus, *Research Notes* 15, 6–8.
- Ingham, K and Thighe, D (2006) Issues with developing a test in LSP: the International Certificate in Financial English, *Research Notes* 25, 5–9.
- Nation, P (2001) Using small corpora to investigate learner needs: Two vocabulary research tools, in Ghadessy, M, Henry, A, Roseberry, R (Eds), 31–45.
- Nation, P and Kyongho, H (1995) Where would general service vocabulary stop and special purposes vocabulary begin?, *System* 23 (1), 35–41.
- Nelson, M (2000) A Corpus-based study of Business English and Business English Teaching Materials, unpublished PhD thesis, University of Manchester.
- Read, J (2000) *Assessing Vocabulary*, Cambridge: Cambridge University Press.
- Scott, M (1999) *WordSmith Tools 3.0*, Oxford: Oxford University Press.

Exploring lexical differences in General English Reading papers

FIONA BARKER RESEARCH AND VALIDATION GROUP

Introduction

This article summarises an exploratory study which used *WordSmith Tools* (version 3; Scott 1999) to provide a lexical analysis of General English Reading papers as a first step towards identifying criterial differences among the Reading texts in question papers at the five different Main Suite levels (KET, PET, FCE, CAE and CPE). It was anticipated that this research would feed into ongoing work to define the construct of Reading, including its context validity. This in turn would inform the volume on Reading currently in preparation for the Studies in Language Testing series which follows that recently published on Writing (Shaw & Weir 2007).

Background

Recent advances in corpus linguistics mean we can explore criterial features of texts, e.g. learners' written or spoken output and exam papers or course materials, more comprehensively than ever before. Corpora and related text analysis techniques play an increasingly important role in describing and codifying linguistic and other features including the lexical, structural and functional aspects of texts. The research reported here provides a lexical analysis of Main Suite Reading papers as part of a larger project to explore the structural, functional and lexical aspects of Reading texts at different proficiency levels. The research aimed to answer the following questions:

1. What lexis is found in Main Suite Reading papers?
2. How does this differ between proficiency levels (A2–C2 on CEFR)?
3. How does this compare to native speaker (NS) general and academic word lists?

These questions were necessarily broad in scope to enable refinement in future studies. Alongside the lexical analysis reported here, we considered the results of previous studies which aimed to establish criterial differences between candidate *output* at different proficiency levels which are summarised below.

Studies of candidate writing

A number of other Cambridge ESOL studies have investigated criterial differences between different levels of Writing. For example, Hawkey and Barker (2004) established three criteria: *sophistication of language*, *organisation and cohesion*, and *accuracy* which they identified for four levels (approximately B1–C2). Their qualitative analyses of learner scripts identified and exemplified features consistently characterising their

proficiency levels. For example, here is the Level 5 from the draft *Common Scale for Writing* (C2: Mastery):

1. Can write extensively and enhance positive impact on the reader through the effective use of sophisticated language resources such as:
 - the ability to vary style of expression and sentence length for effect
 - the use of advanced vocabulary and word order
 - the use of idiom and humour.
2. Can write with only very rare, minor errors of grammar or vocabulary.
3. Can organise extended writing effectively, linking ideas appropriately with or without explicit linking words.

Further evidence was provided by Barker and Taylor (2006) who investigated the impact of American (AmEng) and other varieties of English on learners' written output at FCE, CAE and CPE level. They found:

- A decrease in the proportion of AmEng spelling as proficiency level increases but an increase in the proportion of AmEng lexis as proficiency level increases.
- Learners are more likely to use distinct AmEng lexical items (e.g. *film* and *movie*) than AmEng spelling variants (*colour* and *color*) in the same written text.
- Little evidence for the influence of American (or other) varieties on morpho-syntactic features: all British English variants were more common than other varieties and no relationship was observed between the relative frequency and proficiency level of two specific features (count/non-count nouns and verb phrase collocations).

These findings may also apply to Reading texts and future research could see if this is indeed the case.

The following sections present the methodology used and a summary of the research findings.

Methodology

One Reading paper was selected for each Main Suite level from a recent live administration (the combined Reading/Writing paper was selected for KET and PET).¹ The electronic exam papers were obtained from Cambridge ESOL's item bank (see Marshall 2006). For the two lowest levels – KET and PET – Writing tasks were removed so only the Reading passages remained. All question letters, numbers and rubrics were removed from all levels, keeping only the reading passages and multiple-choice questions as these

1. The papers used are available within past paper packs see www.CambridgeESOL.org

were considered to be input texts for the candidates as much as the Reading passages themselves. Each document was then saved as a text file to allow lexical analysis using WordSmith.

The lexical analysis took place in three stages. Firstly, six word lists were created; one for each exam and one composite list for all five exam papers. Secondly, these were compared using the *WordList* tool to perform a Detailed Consistency analysis within WordSmith to ascertain lexical (non)overlap between the levels of single lexical items. A third stage compared the five resulting Reading text word lists with three other lists, namely the Academic Wordlist (AWL; Coxhead 2000) and the most frequent 1000 and 2000 words in English using *VocabProfile* software within Tom Cobb's online *Compleat Lexical Tutor*. Table 1 shows the text lengths per exam based on the WordSmith token count.

Table 1: Text lengths per exam

Exam	Word count
CPE	4091
CAE	4110
FCE	2845
PET	2104
KET	932

There is clearly an increase in the total amount of material to be read in each paper as the proficiency level of the exam increases, from around 900 to 4000 words. We should note that the overall length of reading texts varies within set limits between different administrations of the exam (as set out in the exam handbooks available online) and that the word counts in this table include reading passages and multiple-choice questions so are likely to be higher than the specifications indicate.² The lexical analysis undertaken is described in the following sections.

Lexical analysis

The *WordList* tool was used to make both a composite word list from all five text files and individual word lists, one per exam. The related Statistics tables for the individual word lists were viewed which showed the Standardised Type Token Ratios (STTR) for each text file. The normal Type Token Ratio is a measure of the proportion of different words in each text and the standardised version is calculated on a sample of words from each text file, in this case 900 words in order to give a rating for KET as this Reading paper contained the fewest number of tokens (see Table 1). Table 2 shows the STTRs for the five exam papers; the higher the number, the larger the proportion of different words in a text or vocabulary range.

2. Note that FCE and CAE Reading specifications have been updated for administrations from December 2008 onwards (see *Research Notes 30* on the FCE and CAE Review Project for details).

Table 2: Standardised type: token ratios

Exam	STTR
CPE	49.89
CAE	45.19
FCE	47.93
PET	42.83
KET	41.22

This table shows an overall increase in STTR as the proficiency level increases indicating a broader vocabulary range as proficiency increases. This finding is in line with expectations, with the exception of FCE/CAE. The next stage in the analysis investigated the overlap (or not) of lexis between the levels.

Lexical overlap between levels

The second stage consisted of running a Detailed Consistency analysis from within the *WordList* tool on all five exam word lists. The output of this analysis shows the frequency of occurrence of all words (single lexical items) and which text files they occur in. This analysis listed 3275 different words which occur at one or more levels with some occurring at all levels (5%), similarly some words at four (4%), three (8%), two (18%) and one level (64%) only (see Table 3 for a breakdown).

Table 3: Number of words occurring at 1 or more levels

Words occurring in <i>n</i> levels	No. words	Percentage (cumulative)
1 level	2094	63.94 (63.94)
2 levels	603	18.41 (82.35)
3 levels	274	8.37 (90.72)
4 levels	133	4.06 (94.78)
5 levels	171	5.22 (100.00)

This analysis shows that a decreasing number of words occur at two or more levels which indicates that the words used in Reading papers at each proficiency level differ. The nature of these words – whether structural or functional – was not investigated further in this study but these findings would be an ideal starting point for another study. Nevertheless, the number of words which occurred at one or more levels was looked at in more detail, as described below.

The list of 2,094 words occurring in one level only was converted to an Excel worksheet in order to ascertain at which level these words occur. This analysis, summarised in Table 4, revealed that an increasing number of words only occurred at higher proficiency levels (FCE, CAE and CPE) and that no words only occurred at PET level.

Table 4: Words only occurring at one level

Words occurring in	No. words
KET only	86
PET only	0
FCE only	458
CAE only	698
CPE only	852

It is not clear why no words only occurred at PET level; this may be due to the nature of the particular PET reading paper analysed and could be checked against other PET papers if this research were replicated. Whilst space does not permit a full discussion of the words only occurring at one level, here are the most frequent examples from each level together with the number of times they occurred in the Reading paper (which could be in one or more texts):

- KET: violin (6), badgers (5), met (3), phone (3)
- FCE: boy (11), robots (6), pages (5), cleaning (3)
- CAE: diary (34), cameras (20), honey (15), bird (14)
- CPE: weather (29), jazz (12), climate (7), advertisements (6).

These words broadly suggest the topics of the Reading texts in the exam papers and from this evidence it is possible to infer that the non-overlap of vocabulary in these Reading texts is partly due to their topic focus as well as their proficiency level. Note that proper nouns are not included in the list above as they are thought unlikely to occur in more than one text unless they are commonly known e.g. *London*.

The next analytical step was to compare the number of words only occurring in two levels, both contiguous (e.g. KET:PET) and non-contiguous levels (e.g. KET:FCE).³ Table 5 shows the partial results of this analysis (note that there were no words just shared by PET and FCE, PET and CAE or PET and CPE).

Table 5: Words only occurring in two levels

Contiguous levels		Non-contiguous levels	
Words occurring in	No. words	Words occurring in	No. words
KET:PET	265	KET:FCE	13
PET:FCE	0	KET:CAE	13
FCE:CAE	95	KET:CPE	20
CAE:CPE	122	PET:CAE	0
		PET:CPE	0
		FCE:CPE	75
Sub-total	482	Sub-total	121

There do not seem to be any general conclusions which can be drawn from Table 5 in terms of a directional increase or decrease in the co-occurrence of single lexical items across levels, although it is interesting to note that there is a greater number of words co-occurring in contiguous than non-contiguous levels (482:121). Whilst there is no discernible pattern between contiguous levels – KET:PET share the largest number of words but PET:FCE share no words – there is an observable increase in the number of shared words as proficiency distance narrows in non-contiguous levels (i.e. KET:CAE – a distance of three levels has fewer shared words than FCE:CPE – a distance of two levels).

When we look at a few examples of shared lexis between contiguous levels we begin to get an insight into the nature of the reading texts at the appropriate levels (the numbers

indicate frequency for the lower then the higher level, e.g. *writers* occurs 16 times in KET at A2 level then 16 times in PET at B1 level):

- KET and PET: *writers* (16/16), *bus* (13/12), *café* (12/12), *wants* (7/7)
- FCE and CAE: *guide* (9/7), *future* (6/3), *money* (6/2), *island* (5/5)
- CAE and CPE: *advertising* (2/16), *clients* (7/4), *product* (1/7), *almost* (3/4).

When we look at shared lexis between non-contiguous levels we note that the overall frequency of the items is lower than in the examples above:

- KET and FCE: *kitchen* (1/2), *useful* (1/2), *boring* (1,1), *clever* (1,1)
- KET and CAE: *brown* (6/1), *badger* (2/1), *keeping* (1/7), *call* (1/2)
- KET and CPE: *play* (6/1); *four* (2/2), *answer* (2/1), *month* (2/1)
- FCE and CPE: *blue* (1/17), *real* (1/6), *learnt* (2/3), *smile* (1/4).

This analysis of lexical overlap between levels generally shows that a larger number of words occur at only one or two levels than three or more, as one might expect given the interaction of proficiency level and topic specificity on vocabulary in Reading texts. At this stage, however, the nature of these words (whether they are content or grammatical words, for example), was not taken into account which could alter the findings somewhat. In order to begin to address this, the third stage of the lexical analysis took place, namely a comparison of the Reading text word lists with other widely recognised NS word lists.

Comparison of Reading text word lists with NS word lists

The individual Reading text word lists were run against the Academic Wordlist (AWL; Coxhead 2000) and the most frequent 1000 and 2000 words in English (versions of West's 1934 General Service List, see <http://jbauman.com/aboutgsl.html>) using the online VocabProfile software on the Compleat Lexical Tutor website. This was done to check the proportions of general versus academic vocabulary in the Reading texts. The five Reading text word lists were pasted in to the software and a display showed each text colour-coded according to the vocabulary within it, i.e. whether it was from the first 1000 words, the second 1000 words, the AWL or not from any list. According to the VocabProfile website, 'A typical NS [native speaker] result is 70–10–10–10, or 70% from first 1000, 10% from second thousand, 10% academic, and 10% less frequent words'.⁴ Table 6 shows the percentage of words within each level's Reading paper which match the first 1000 words in English (K1), the second 1000 words in English (K2), both K1 and K2 together, the Academic Wordlist (AWL) or none of the three (off-list), compared with the NS average.

3. See www.CambridgeESOL.org for an indication of the CEFR levels of our Main Suite and other exams.

4. See <http://www.lextutor.ca/vp/research.html>

Table 6: Reading text words in 1st 1000, 2nd 1000, Academic Wordlist and off-list

Exam	K1 words	K2 words	K1+K2 words	AWL words	Off-list words
CPE	75.02	7.65	82.67	4.06	13.26
CAE	78.98	6.40	85.38	4.69	9.92
FCE	81.36	7.91	89.27	3.50	7.24
PET	83.33	6.76	90.09	2.76	7.14
KET	85.55	5.89	91.44	0.21	8.35
NS average	70.00	10.00	80.00	10.00	10.00

We can observe a number of interesting findings from Table 6. Firstly, none of the Reading papers displays the 'typical' NS pattern in terms of the proportion of words from each of the five categories. We can say, however, that the nearest to NS in terms of K1 words is CPE, the nearest to NS in terms of K2 words is FCE followed closely by CPE, the nearest to NS in terms of K1+K2 words is CPE, whereas CAE is the nearest to NS in terms of AWL and off-list words. As these findings concern the three higher levels of Main Suite (FCE-CPE, B2–C2) this is still somewhat encouraging. Looking in more detail at the results, there is a preference for general vocabulary over academic vocabulary in the Reading texts which is as expected as these are General English exams. If an academically focused reading paper were compared (e.g. IELTS Academic Reading), we would expect the proportions to be slightly higher in favour of the academic vocabulary although this comparison was not done within this study. If a Business or Financial English oriented text were used (e.g. see Wright 2008) we would expect the proportions of general and academic vocabulary to be lower with more off-list words. This might reveal possible contenders for a core financial or business vocabulary (cf. Wright 2008) and the proportion of vocabulary across different genres would also be informative.

Returning to Table 6, there is a decrease in the proportion of K1 words as the proficiency level increases (86–75%) which is as expected, as general vocabulary gives way to more specialised and/or rarer vocabulary at the higher proficiency levels. There is a smaller increase in the proportion of K2 words as the proficiency level increases (6–8%) which is a somewhat surprising result for the Upper Main Suite, especially CAE, which needs to be explored further. However, Table 6 shows a small decrease in the proportion of K1 and K2 words combined as the proficiency level increases (91–83%) which again is as expected. In terms of academic vocabulary, there is an increase in the proportion of academic words as the proficiency level increases (0.2 to 5%) which is generally as expected although the CAE text had a higher proportion than the CPE text in this sample which may be due to the texts investigated rather than a feature of CAE Reading texts in general. Finally there is an increase in the proportion of off-list words as the proficiency level increases (7–13%) which is as expected, although the higher proportion for KET over PET and FCE may be due to a task effect in the Reading papers selected for this study. Alternatively, it may be due to the A2 level of vocabulary contained in the KET Reading

texts being of a lower level than the 1st and 2nd 1000 words in English which may not include certain vocabulary which appears at KET level. This could be investigated by comparing the existing KET word list with the K1 and K2 lists, another idea for a follow-up study.

Findings

This analysis suggested the following general criterial features for distinguishing Reading texts in Main Suite exams in terms of their lexical features:

- Vocabulary range: the standardised type: token ratio increases as the level increases (more tokens and more types = broader vocabulary range)
- Overlap between levels: some evidence for a larger number of words occurring at only one or two levels as the proficiency level increases
- Proportion of general versus other types of vocabulary: there is a greater proportion of words from academic English and a lower proportion of general English words as the proficiency level increases.

Returning to the research questions, this research established that, in the sample of Reading texts investigated, a broad range of lexis was found in Main Suite Reading papers (over 3200 different types in 14,000 tokens, an average STTR of 45.41 in a range of 41 to 50). The range and type of lexis differed between proficiency levels in a number of ways. The study of the overlap of lexical items showed that there is a decreasing number of words occurring at two or more levels although 36% occurred at two or more levels in this sample. Of this 36%, the analysis of words occurring in two levels only show that more words co-occur at contiguous than non-contiguous levels which suggests that lexis in Reading papers in Cambridge ESOL exams is carefully selected and applied. It should be noted, however, that this finding may not be borne out by further analysis of the words occurring at three or more levels.

The third research question considered Reading text lexis versus native speaker lexis (both general and academic). The findings revealed atypical distributions of general and academic vocabulary versus NS usage although higher proficiency level texts (FCE, CAE and CPE) were closest to NS usage in terms of the proportion of words from each of the NS word lists analysed. Generally, findings in this area were as expected for general and academic usage given the nature of the General English Reading texts studied, with the exception of KET which merits further study.

The findings provided by this study indicate that further research is needed on these Reading texts – and others – in order to explore these and other areas.

Conclusion

This study has provided informative findings about the lexis contained in Main Suite Reading exam papers and has shown that this differs in a number of ways according to proficiency level. A methodology for using WordSmith Tools to explore Reading texts lexically has been established and this work has highlighted some criterial differences among the Reading texts in papers at the five different Main Suite

levels in terms of their lexis. The criterial differences outlined here could therefore feed into our understanding of the construct of Reading in General English exams and the context validity of Reading.

As with any study further work is recommended, including a replication of the WordSmith analysis using more text files, as only one Reading paper per level may not have provided generalisable results although this should be the case as all exam papers are subject to a rigorous and lengthy production and pre-testing process (see Green and Jay 2005). Additionally, the lexis which occurred in three or more levels (578 words) could be explored further, as this would, for example, reveal lexis that is shared by PET and other levels thus revealing more than the co-occurrence in two levels reported here. Further analysis could also categorise the lexis found in each level according to an existing classification (see Hughes 2008) or comparing it to candidate output taken from the Cambridge Learner Corpus, our unique collection of 30 million words of candidates' written exam scripts.

Going beyond the lexical analysis reported here, further work could be undertaken to analyse the criterial features in Reading texts along structural and functional lines. In order to do this, text files could be tagged and parsed to analyse linguistic structures at the different levels and the structures listed in current exam handbooks could then be explored in the Reading texts. Additionally, a list of functions could be drawn up from handbooks and explored in the Reading tasks, possibly using a checklist as was developed for Speaking Tests (see Brooks 2003). These analyses would enable the following hypotheses to be explored:

- Structural range and complexity increases as the level increases.
- Specific structures are only found at certain levels.
- Functional range and complexity increases as the level increases.
- Specific functions are only found at certain levels.

Work is underway within *English Profile* to look at these areas for Writing and, to a lesser extent, for Reading (in the

analysis of course books and other teaching materials). The English Profile research team is working towards Reference Level Descriptors for English which will provide a uniquely detailed and objective analysis of what levels of achievement in language learning actually mean in terms of the grammar, vocabulary and discourse features that learners can be expected to have mastered at each proficiency level of the CEFR. The research described here will form part of our contribution to the English Profile endeavour.

References and further reading

- Academic Wordlist
<http://language.massey.ac.nz/staff/awl/awlinfo.shtml>
- ALTE Can-do statements www.alte.org/can_do/index.php
- Barker, F and Taylor, L (2006) *Using corpus studies to explore the impact of regional varieties of English on learners' written performance*, paper presented at BAAL/IRAAL, Cork, 7–9 September 2006.
- Brooks, L (2003) Converting an observation checklist for use with the IELTS speaking test, *Research Notes* 11, 20–21.
- Compleat Lexical Tutor www.lextutor.ca
- Coxhead, A (2000) A New Academic Wordlist, *TESOL Quarterly* 34 (2), 213–238.
- English Profile www.englishprofile.org
- Green, T and Jay, D (2005) Quality Assurance and Quality Control: Reviewing and pretesting examination material at Cambridge ESOL, *Research Notes* 21, 5–7.
- Hawkey, R and Barker, F (2004) Developing a Common Scale for the Assessment of Writing, *Assessing Writing* 9 (2), 122–159.
- Hughes, G (2008) Text organisation features in an FCE Reading gapped sentence task, *Research Notes* 31, 26–31.
- Marshall, H (2006), The Cambridge ESOL Item Banking System, *Research Notes* 23, 3–5.
- Scott, M (1999) *WordSmith Tools Version 3.0*, Oxford: Oxford University Press.
- Shaw, S D and Weir, C J (2007) *Examining Writing: Research and practice in assessing second language writing*, Studies in Language Testing volume 26, Cambridge: Cambridge University Press/UCLES.
- Wright, A (2008) A corpus-informed study of specificity in Financial English: the case of ICFE Reading, *Research Notes* 31, 16–21.

Text organisation features in an FCE Reading gapped sentence task

GLYN HUGHES ASSESSMENT AND OPERATIONS GROUP

Introduction

This article is based on research into the textual organisation of a Reading task undertaken as part of a Masters degree in TESOL at Aston University (UK). The study uses a range of methods to analyse the text of a 'gapped sentence' task used in the Cambridge First Certificate in English (FCE) Reading test, which aims to test candidates' knowledge of text structure. The text used in the examination is compared to the newspaper article that it

was based on with a view to discovering more about two areas:

- How the discourse features of the adapted text compare to the discourse features of the original text and the extent to which this affects the manner in which candidates complete the task and therefore the validity of what is being tested.
- What discourse features candidates are able to use to answer the questions correctly and whether the

sentences removed from the text create a valid test of candidates' ability to recognise discourse features.

Research such as this adds to Cambridge ESOL's knowledge about its tasks and can be used when training item writers and when editing new material.

The gapped sentence task

The gapped sentence task aims to test 'text structure, cohesion and coherence' and consists of 'a text from which sentences have been removed and placed in a jumbled order after the text' (Cambridge ESOL 2007:7). Candidates have to decide from where in the text the sentences have been removed. Each gapped sentence text is adapted from an original text to reduce its level of linguistic complexity and also to produce a workable task that does not have more than one correct answer for a question. In order to genuinely test what it claims to be testing, it is essential that, in spite of adaptations made, the features of 'text structure, coherence and cohesion' in the text reflect texts found in the real world. Too many artificial changes that disrupt normal organisation of discourse will make it a less effective task. In addition, it is essential that the sentences are removed in a way that enables candidates to use features of text organisation to complete the task.

The text selected for analysis in this study is based on a 'human interest' newspaper feature article about Linda Greenlaw, a female swordfish-boat captain who became famous after she guided her crew safely through a major storm at sea. The article has a clear narrative thread and is also trying to make a wider point – that of a woman fighting prejudice and becoming a hero – and in making such a point focuses on specific episodes. Articles such as this appear frequently in the FCE Reading test, especially in the gapped sentence task, as a non-chronological narrative is helpful when constructing the tasks. Therefore, one reason for choosing this text is that item writers deal with many similar texts and should be able to transfer any findings from this research to texts they are currently working on. Additionally, the selected text has been significantly amended for use in the exam. Whilst the amount of amendment is more than usual for an FCE reading text, more can be learned about the effects of adapting texts from one with a large number of amendments.

Methodology

A range of methods, outlined below, is used to analyse the texts as candidates can utilise a range of different types of textual knowledge (both conscious and subconscious) in order to complete the task. The analysis first looks at the text-level patterns of thematic organisation, patterns and signals and cohesion before analysing each paragraph of the text in detail.

Thematic organisation

The analysis of thematic organisation follows the model developed by Fries (1994:229–249) based on Halliday's principles of thematic organisation (see Bloor and Bloor

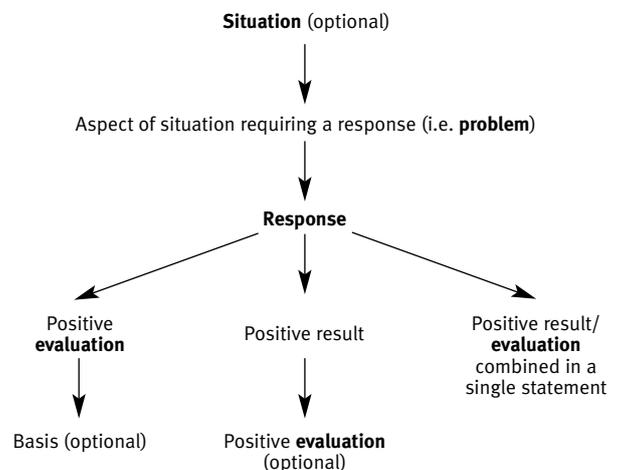
1994:65–85). Fries argues that 'writers use position at the end of a clause to indicate the newsworthy information to their readers and ... use the beginning of a clause to orient their readers to the message which will come in the rest of the clause' (1994:233–4). Fries also develops the concept of the N-Rheme: the last, and newsworthy, part of the clause and further argues that the goals of a text are largely found in the N-Rhemes, as this is the part of the message that is emphasised (ibid:236).

For a newspaper article, the pattern of an N-Rheme giving newsworthy information should be readily apparent. Therefore, the original text is analysed for this pattern and then compared to the adapted text. Following Fries' example, and to lend coherence to the analysis, the thematic organisation of rankshifted clauses is not analysed as these rankshifted clauses, especially when postmodifying a noun at the end of a clause, often form part of the N-Rheme, as is the case with 'where families tortured by anxiety lined the dock' analysed below.

Patterns and signals

The basis for the analysis of patterns and signals in the text below is the work of Hoey (1994,2001) and Winter (1994). Hoey (2001:119–141) argues that the most culturally popular form of organisation of written text is a Situation, Problem, Response, Evaluation (SPRE) pattern, which he outlines as shown in Figure 1.

Figure 1 : Hoey's SPRE pattern (adapted from Hoey 2001:127)



Hoey (1994:44) states that 'the English language indicates to the reader/listener the functions of a particular discourse's constituent sentences' through *signalling* and that 'problems of comprehension can be shown to arise from "faulty" or missing signalling'. Hoey uses Winter's definition of Vocabulary 1 (subordination), Vocabulary 2 (closed system sentence connectors) and Vocabulary 3 (lexical items) as a basis for analysing a sample text (ibid:26–45).

In discussing *clause relations*, Winter argues that because a writer will always try to ensure that each clause he writes is relevant, 'we have got to assume that relations between the clause in its sentence and its adjoining sentences cannot be random or haphazard' (1994: 48).

He describes two kinds of clause relations (ibid:50–53):

- *matching relations*: generalisation-example, preview-detail, matching-compatible, matching-contrast, hypothetical-real, denial-correction
- *logical sequence relations*: cause-consequence, achievement-instrument.

This analysis will use Winter's categorisations to gain insight into how candidates can complete the gapped sentence task by looking at the relations between clauses and how they are signalled.

Cohesion

In addition to the above features, the surface level clues that candidates can use to complete the task are also considered. These include referencing within the text, possible uses of ellipsis and substitution, lexical cohesion and conjunctive cohesion (Halliday 1994:334) and also what Francis refers to as 'advance' and 'retrospective labels' (1994: 83–101).

We now turn to an analysis of the original and amended versions of the text under investigation.

Text-level analysis

As is typical of this kind of human-interest article, there is a clear pattern of Situation, Problem, Response and Evaluation:

Situation: Linda Greenlaw is in love with the sea and wants to become a fisherman.

Problem: Fishing is a bastion of male prejudice.

Response: Greenlaw works extra hard to become a captain and to earn the men's respect.

Evaluation: She succeeds as evidenced by saving her crew in a storm and her continuing success as a captain.

In addition to this main pattern, there are many other SPRE patterns running through the text, as discussed below in the paragraph-by-paragraph analysis.

Unsurprisingly, *Linda* is the theme in the vast majority of clauses in the original text. The only paragraphs where the theme is not usually Linda are paragraph A, setting the scene, and paragraph F, which focuses on a book about Linda written by Sebastian Junger. Broadly speaking, the adapted text follows a similar pattern, with the exception that the vignette about the book is less detailed, taking up only half a paragraph.

The N-Rhemes in both the original and the adapted text broadly fall in line with the SPRE patterns; for example paragraph B, which outlines Linda's problems, contains a large amount of negative vocabulary in the N-Rhemes. Towards the end of the text, when Linda's success is being discussed, the N-Rhemes are generally much more positive, although this is more evident in the original text, which finishes with a very positive evaluation of life at sea.

Paragraph-by-paragraph analysis

There now follows a detailed analysis of selected paragraphs from the adapted text (B,C,D,F,G,H), with

reference to the original. The aim is to explore some examples of how the text-level patterns mentioned above are manifested in the text. In the paragraphs below the text is divided into non-rankshifted clauses that are labelled (B3a, B3b etc) according to paragraph (B), sentence (3) and clause (b). Within each clause, the Theme is shown in bold, the N-Rheme is italicised and the highlighted sentence is the gapped sentence which was tested; candidates have to work out where each should be inserted within the text.

Paragraph B original text

B3a **Clever and determined, Linda Greenlaw** had *horrified her middle-class parents*

B3b **when she** *abandoned plans to go to law school*

B3c and became a professional fisherman facing *danger – and prejudice – in a tough male world.*

B4a **But she** *silenced her critics forever after the deadly storm*

B4b **when she** *guided the 100 ft Hannah Boden safely back to the harbour, where families tortured by anxiety lined the dock.*

Paragraph B adapted text

B4a **Clever and determined, Linda Greenlaw** had *horrified her middle-class parents*

B4b **when she** *abandoned plans to go to law school to become a professional fisherman, facing danger – and prejudice – in a tough, male world.*

B5a **But she** *silenced her critics forever after that deadly storm,*

B5b **when she** *guided her boat safely back to harbour.*

B6 **Anxious families** were *lining the dock* there when it arrived.

Paragraph B in the adapted text contains the response (B5b) and evaluation (B5a) to a problem outlined in Paragraph A (a dangerous storm). Before doing so, however, it also raises a second problem (B4), which is also responded to and evaluated in the same paragraph. In this sense, the gapped sentence (B5) has been clearly set up for the candidates across two paragraphs. The SPRE pattern is signalled in many ways. Firstly, Vocabulary 2 *but* (B5a) sets up a positive evaluation from a negative situation in the form of adversative conjunctive cohesion.¹ Then, Vocabulary 3 *safely* (B5b) positively evaluates the response – *guided her boat back to harbour* – required of Linda in the situation outlined in A. Also, in B5a, *silenced* and *forever* are strong positive evaluations of her response: guiding the boat back safely. Additionally, *critics* is a superordinate for her horrified middle class parents and the people from whom she faced prejudice. For these reasons this is a fair test of 'text structure, cohesion and coherence'.

1. *But* is usually labelled as Vocabulary 1 according to Hoey and Winter's classification. However, as is standard in journalistic text, it is used in this text as a sentence connector so in these instances I have labelled them Vocabulary 2.

There are, however, some consequences of amending the text that are perhaps not ideal. The rankshifted clause *where families tortured by anxiety lined the dock* in B4b of the original text is turned into a separate sentence in the adapted text (as also happened in paragraph A). This was done to create a workable test item through lexical cohesion: *it* is a pronominal reference to *boat*, and *there* is a reference to *to harbour*. However, it also has the effect of creating the only theme that isn't Linda. It also turns the negative *tortured by anxiety*, clearly emphasised in the N-Rheme, into a rather inappropriate theme. Finally, this last sentence is rather superfluous to the SPRE pattern of the first two paragraphs. It is providing detail, but only about the *harbour*, which is the nominal head of the indirect complement of the verb *guided* in the previous clause.

Paragraphs C and D original text

- C5a **She** had *not slept for 36 hours*
 C5b **as she** manned the helm in *a heroic fight for survival*.
 C6 **But the six-man crew of another swordfishing boat**, the Andrea Gail, were *all lost in the depths off Newfoundland with their experienced skipper, her good friend Billy Tyne*.
 D7 **As one of Linda's crewmen, Wally Tallaksen**, puts it:
 D8 **'She** earned her stripes.
 D9 **It didn't matter if Linda had been man, woman or dog**.
 D10 **She** got us through it.
 D11a **From then on, she** was never "Linda",
 D11b **she** was always "Cap'n".'

Paragraph C adapted text

- C7a **She** had *not slept for 36 hours*
 C7b **as she** led the crew in *a heroic fight for survival*.
 C8 **As one of Linda's crewmen, Wally Tallaksen**, puts it:
 C9 **'It didn't matter that Linda wasn't a man**.
 C10 **All we cared about** was *survival*.
 C11 **She** got us through it.
 C12a **From then on, she** was never Linda,
 C12b **she** was always Captain.
 C13 **That** showed our respect for her.'

Paragraph C is a composite of paragraphs C and D in the original. Although the original paragraphs have similar themes, the N-Rhemes give the paragraphs rather different messages. Paragraph C focuses on the difficulty with *not slept for 26 hours, heroic fight for survival*, and the story of the crew that died on another boat. Paragraph D on the other hand focuses on the respect that Linda earned with phrases such as *earned her stripes, got us through it* and the contrastive pattern of *never Linda, always Captain*. The adapted text combines the two paragraphs but because it removes the paragraph about the dead crew, successfully

focuses the message on Linda's heroics. This is an example of successful text adaptation.

The gapped sentence C12 has a logical-sequence relation to the sentence before it as it is describing a consequence – *she was always captain* – of what happened before – *she got us through it*. It also gives Wally's evaluation of Linda's response (C11) to the problem (C7–10). However these relations are only signalled implicitly. The temporal conjunctive cohesion of *from then on* signals that C12 has happened ever since C11, implying causality. Further, there is the implied positive of being called *Captain* rather than *Linda*. The implicit nature of the evaluation could be argued to help candidates as, unlike the evaluative gapped sentence in paragraph B, the fact that there is no need for any adversative conjunctive cohesion means that it is likely to follow a positive comment.

Although only implicit at this stage, the exam text adds C13, which makes the cause-consequence relationship explicit. As in paragraph B, the sentence after C12 helps to tie the gapped sentence in cohesively. *That* is a demonstrative reference to the crewmen's reaction. Candidates can also make use of the pronominal reference *she* in C12a and b, which refer back to Linda in C9. This is further evidence of the importance attached to cohesive devices when editing this text.

Paragraph D adapted text

- D14 **Linda's feat of seafaring** made her *an instant celebrity*.
 D15 **And now, eight years after the storm, she** has become *a worldwide publishing phenomenon*.
 D16a **The New York Times** has *predicted* that
 D16b **The Hungry Ocean, her autobiographical account of how she broke into a world dominated by men, will sail onto the bestseller list**
 D16c **when it comes out**.
 D17 **However, completing it wasn't something that came naturally to her**.
 D18 **She** explains *why*:
 D19 **'Being landbound writing this book was the toughest thing I've ever done**.
 D20 **I** longed to go back out there, *with the wind in my hair and the sea spraying onto my face*.
 D21 **The sea** is *dangerous and unforgiving*.
 D22 **But it** is also *majestic, spiritual, calming and powerful*.
 D23 **I** love it to *the bottom of my soul*.
 D24 **Out on the ocean**, you *feel free*.
 D25 **And even storms** can *be beautiful*.'

Paragraph D in the adapted text is a composite of three different parts of the original text. The theme of the clauses changes all the way through the first half of the paragraph. The N-Rheme also shifts in the paragraph: in D14, 15 and 16b, the focus is on positive evaluations of the impact on Linda's career of that night – *instant celebrity, worldwide publishing phenomenon, will sail onto the bestseller list*.

Sentences D17 and D19 both have processes as their theme, in addition to D17's textual theme: *however*. They both have N-Rhemes giving a negative evaluation of the process. In the last few clauses of the paragraph, D21, 22, 24 and 25 are all about the sea and the N-Rhemes are evaluating it, mostly positively. Although it does not parallel the original, this paragraph could be argued to hang together as a transitional paragraph from the first section of the article about the storm to the second section about Linda becoming a fisherman.

The SPRE pattern in paragraph D is interesting in that it sets up a situation (Linda wrote a book) and a problem (she hated being away from the sea), but then does not provide a response or an evaluation. Although this part of the text has been adapted significantly, the same is true of the original article. In spite of this, candidates are able to use clause relations to place the gapped sentence. D16 and D17–25 have a matching-contrast relationship, as signalled by the Vocabulary 2, *However* at the beginning of D17. Additionally, D17 is the preview for the detail expressed in the rest of the paragraph, signalled clearly lexically by evaluative phrases such as *toughest thing I've ever done* (D19) and *I longed to go back out there* (D20). As in paragraph C, the clause relations are signalled more explicitly in the adapted text than in the original, with the addition of *she explains why* in D18. Candidates are given further cohesive clues from *it* (D17) – something that needed completing – referring to *The Hungry Ocean* in D16.

Paragraph F adapted text

F32a **She worked her way through the ranks,**

F32b *toiling as a cook, cleaner and deck-hand*

F32c *before being made a captain at the age of 24.*

F33a **I** came back into port *that first day in charge*

F33b **and [I]** *landed the ship badly.*

F34 **Basically,** *I crashed into the dock!*

F35 **The jokes about "women drivers"** *haunted me for years!*

The gapped sentence in this paragraph (F32) was the hardest in the test when it was administered. In terms of the SPRE patterns, F32a and b form part of the response, together with E31b *so I went*, to the problem that arises in paragraph E – *When I graduated, I knew I could never be stuck in an office. I could hear the ocean calling me*. In fact, there is a preview-detail relationship between E31b and F32a and b. F32c is a positive evaluation of Linda's work in F32a and b. F33–35 then introduce a new problem. What may make this difficult is a lack of explicit signalling to indicate this problem arising. In particular there is no conjunctive cohesive device to signal the shift from positive outcome – being made captain at 24 – to the negative – crashing into the dock. In fact, the only explicit link between F32 and F33 is *that first day in charge* referring back to *being made a captain*. However, because this link exists and because the gapped sentence fits well in terms of text structure, although the question was difficult, it is fair, as was demonstrated by the fact that it discriminated well between strong and weak candidates.

Paragraph G adapted text

G36a **Such comments, naturally,** *infuriated her*

G36b **but they also** increased her determination to prove herself *in the most masculine of professions.*

G37a **Linda** earned the men's admiration *the hard way:*

G37b working *longer hours*

G37c **and** *pushing herself to the brink, physically and mentally.*

This paragraph contains the response and evaluation to the problems raised in paragraph F. The link to the previous paragraph is made clear in many ways. Firstly, *such comments* is an example of retrospective labelling, what Francis calls an 'illocutionary noun' (1994:90). In this case *comments* encapsulates the phrase *jokes about women drivers*.

There are also cause-consequence patterns linking the clauses. The jokes about women drivers had two consequences, signalled in G36 by the clearly evaluative Vocabulary 3 *infuriated her*, and by *her determination to prove herself*. G37 is also a consequence of G36. Linda's *pushing herself to the brink* was a consequence of her *increased determination*. The lexical link here is between *prove herself* and *earned admiration*, and *most masculine of professions* and *men's admiration*.

Paragraph H adapted text

H38a **During a typical 30-day, 2,000-km swordfishing expedition,** *the crew live cramped together,*

H38b hunting fish which can *weigh as much as 250kg each.*

H39a **Affectionately known as 'Ma' in the fleet,** she admits that

H39b **she** can be *a tough, unforgiving boss.*

H40 **But I** *love my boys.*

H41 **I** can *honestly say*

H42 **I don't think** *being a woman has ever worked against me.*

H43a **Guys** *get competitive*

H43b **when they** *see me working.*

H44 **They** *don't want to be beaten by a woman.'*

The gapped sentence, H39, has a matching-contrast relationship with H40, clearly signalled with an adversative *but*. The lexical contrast is provided by *tough, unforgiving* contrasted with *love*. *Boss* and *my boys* add more cohesion with their antonymous relationship. In the original, the gapped sentence (with some modifications to avoid taboo areas) is T47 *Affectionately known as 'Ma' in the fleet (and less endearingly as Moby Dickless), she admits....* The previous sentence – *Linda would regularly gross £10,000 to £15,000 (S46)* – starts with *Linda*, giving a relatively close point of reference for *she* in T47. A weakness of the adaptation is that the *she* refers to Linda three sentences back and in a quite different part of the text. However, as Linda is the only woman in the entire text, it is unlikely to have caused the candidates many difficulties.

Conclusion

Returning to the two areas under consideration, this study has revealed a number of strengths of the gapped sentence task. Firstly, with a couple of exceptions, the exam text maintains the thematic organisation of the original in a way that enables candidates to make sense of the text and complete the task. Secondly, the adapted text has clear patterns and signals that candidates can use. The structure of the text often gives compelling reasons for candidates to place a particular sentence in a particular gap. This structure is appropriately signalled in a way that helps facilitate task completion. Finally, candidates can make use of a range of cohesive devices to complete the task. The validity of what is being tested in the task under investigation is therefore supported, and this gapped sentence task can be viewed as a valid test of candidates' ability to recognise discourse features.

The analysis also reveals a number of issues that would benefit from further investigation. Occasionally, when a sentence is adapted, this has a negative impact on thematic organisation and/or text structure (see paragraphs A and B). Additionally, throughout the adapted text, the gapped sentences are always linked in very strongly using cohesive devices (particularly pronominal referencing), even though this is not always the case in the original text. This may suggest a possible bias towards cohesion over other forms of text organisation when adapting texts for use in the gapped sentence task. These issues can be explored further by examining other texts and, if necessary, addressed in item writer training and in editing meetings.

References and further reading

- Bloor, T and Bloor, M (1995) *The Functional Analysis of English: A Hallidayan Approach*, London: Arnold.
- Cambridge ESOL (2007) *First Certificate in English: Handbook for teachers*, Cambridge: UCLES.
- Coulthard, M (Ed.) (1994) *Advances in Written Text Analysis*, London: Routledge.
- Edge, J and Richards, K (1998) 'May I see your warrant please?' Justifying outcomes in qualitative research, *Applied Linguistics* 19(3), 334–356.
- Francis, G (1994) Labelling Discourse: an Aspect of Nominal-group Lexical Cohesion, in Coulthard, M (Ed.), 83–101.
- Fries, P (1994) On Theme, Rheme and Discourse Goals, in Coulthard, M (Ed.), 229–249.
- Hoey, M (1994) Signalling in Discourse: a Functional Analysis of a Common Discourse Pattern in Written and Spoken English, in Coulthard, M (Ed.), 26–45.
- (2001) *Textual Interaction: an Introduction to Written Discourse Analysis*, London: Routledge.
- Sinclair, J (2004) *Trust the Text: Language Corpus and Discourse*, London: Routledge.
- Weir, C J (2005) *Language Testing and Validation: An Evidence-Based Approach*, Basingstoke: Palgrave Macmillan.
- Winter, E (1994) Clause Relations as Information Structure: Two Basic Text Structures in English, in Coulthard, M (Ed.), 46–68.

Conference reports

Cambridge ESOL staff have recently attended and presented at various national and international conferences and other events organised by the IELTS Partners, ALTE and Council of Europe amongst others, a selection of which are reported on here.¹

IELTS Conference Series 2007

Research and Validation Group staff recently participated in a series of conferences with the British Council IELTS team, designed to build greater understanding about IELTS bandscores and other informational resources for stakeholders. The first conference of the series was the FE Colleges and the International Market 2nd Annual Conference held at Jurys Great Russell Street Hotel, London on 16th October, 2007. Sacha DeVelle (Cambridge ESOL) presented on the standard setting features of the *IELTS Scores Explained* DVD (for the UK market) and the processes involved in deciding which IELTS score a candidate needs

before entering a particular institution or course.

The second event of the series was the European Council of International Schools (ECIS) Annual Conference (21–25 November, 2007), held this year at the Principe Felipe Congress Centre in Madrid, Spain. There were over two hundred guest speakers during the five day event and a major exhibition by school suppliers and publishers. The event attracted approximately 3000 participants who were teachers, heads of school, guidance counsellors and college admissions personnel. Helga Stellmacher (British Council) and Sacha DeVelle (Cambridge ESOL) jointly presented on the features and benefits of the IELTS test for international schools, with a particular focus on the rapidly expanding IELTS market in the US.

The third conference to round off the series was the Education UK Partnership Development Conference (4–5 December, 2007) held at the Edinburgh International Conference Centre, Scotland. Here the focus switched to emerging international markets together with specific country sessions on China, Vietnam, Nigeria and Pakistan. Saima Satti (British Council) presented on IELTS test security and Sacha DeVelle (Cambridge ESOL) demonstrated

1. Thanks to Andy Chamberlain, Margaret Cooze, Michael Corrigan Sacha DeVelle, Angeliki Salamoura and Ivana Vidakovic for their contributions.

practical tasks from the *IELTS Scores Explained* DVD to help stakeholders understand and determine bandscores 5 to 7. The IELTS session attracted over 70 participants from UK institutions and demonstrated a very successful end to the series.

For further information on IELTS please visit www.ielts.org

ALTE language testing courses, September 2007, Valencia, Spain

In September 2007 ALTE ran two very successful courses on language testing at the Universidad Politécnica de Valencia in Spain. The introductory course in Language Testing was taught by Professor Cyril Weir from the University of Bedfordshire, and Professor Barry O'Sullivan from the University of Roehampton. The course in testing Speaking was taught by Dr Lynda Taylor (Cambridge ESOL consultant). The ALTE members who attended represented a wide range of languages including Basque, Catalan, English, Greek, Irish, Italian, Polish, Serbian, Spanish, Turkish and Welsh. Both courses ran over 5 days (24–28 September) and covered a number of topics related to test design, production and validation. Even so, there was still time to enjoy the balmy weather, el Casco Antiguo (the Old City), Mediterranean cuisine and Valencian night life. The overall consensus from participants was extremely positive in terms of course delivery, learning objectives and choice of venue. The course is open to ALTE members, affiliates and others with a professional interest in language testing.

Details of the date and venue for the ALTE Testing course 2008 can be found on the ALTE website: www.alte.org

Council of Europe Seminar on the use of the Manual for Relating Language Examinations to the CEFR 2004–2007, December 2007, Cambridge, UK

The Common European Framework of Reference for Languages: Learning, Teaching, Assessment, (CEFR) published by the Council of Europe, in its own words, 'provides a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe' (Council of Europe 2001:1). By relating or aligning their examinations to the CEFR, examination providers enable themselves and others to more easily compare their exams with those of other providers in other languages. In the long-term, the CEFR is expected to have a positive impact on life in Europe, by improving communication and mutual understanding in Europe, enhancing the practicality of linguistic and cultural diversity in Europe by facilitating comparisons and increasing the mobility of European citizens. Reaping these benefits, however, first requires a great deal of work. In the case of exam providers, the process of relating examinations to the CEFR is complex and time-consuming and the ways in which this is done may vary greatly, depending, in large part, on factors such as the purpose, context and construct of the exam. To help examination providers with this work, the Council of Europe produced a preliminary pilot version of a *Manual for Relating Language Examinations to the CEFR*,

with the intention of revising this edition after collecting feedback from those having used it. Primarily, this seminar was conceived as one means to gather feedback from users of the *Manual* and relay it to the Council of Europe and the authors of the *Manual*. However, it was also thought important to try to make the seminar as useful to everyone as possible, so the seminar aimed to:

- Provide an opportunity to share experiences about the *Manual* in its draft form for those who have actively used it in their own contexts.
- Learn about the main lines of revision in progress based on feedback received and to provide a final general review prior to publication (foreseen 2008).
- Help participants in finalising their own case studies in order to write them up with a view to possible publication at a later stage.
- Provide the Council of Europe with input which may lead to further guidance on appropriate uses of the *Manual* in a variety of contexts.

The seminar took place over two days, was organised by ALTE, hosted by Cambridge ESOL and additional sponsorship was provided by the Goethe-Institut and Test DaF, both from Germany. The seminar began with welcomes from Dr Mike Milanovic, Manager of ALTE and Chief Executive of Cambridge ESOL and Johanna Panthier of the Council of Europe. Eleven presentations followed over the two days, reporting on the work of organisations in nine European countries. The presentations by Cambridge ESOL were *Reflections on using the Draft Manual: the Cambridge ESOL FCE case study*, given by Dr Hanan Khalifa, Dr Angeliki Salamoura and Angela French and *CEFR alignment using the Preliminary Pilot Manual – Asset Languages*, given by Dr Neil Jones. Presentations were interspersed with periods of discussion on issues referred to in the presentations; the whole event was moderated by Dr Lynda Taylor. A presentation slot was also reserved for the authors of the *Manual*, who were present: Dr Brian North, Dr Norman Verhelst and Professor Sauli Takala. The authors spoke about the progress of the revision and their current plans for the new version of the *Manual*.

The event provided ample food for discussion and many important issues were raised – several of which were revisited throughout the seminar presentations and discussions. Among the issues discussed were:

- the way in which cultural influences can affect the interpretation of the CEFR descriptors
- the need for more illustrative samples which mark the boundaries of each common reference level (A1 to C2)
- the need for more supporting evidence for these samples, additional procedures in the *Manual*, greater reference to supporting materials and a bibliography
- the differing needs and resources of large and small examination providers
- the need for a thorough familiarisation phase
- the importance of team building and team balance on an alignment project.

These areas will be considered in the ongoing work to finalise and publish the *Manual*. For further information on CEFR see

www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf
and for information on the Manual see
www.coe.int/T/DG4/Portfolio/documents/Manual%20for%20relating%20Language%20Examinations%20to%20the%20CEF.pdf

34th ALTE meeting, November 2007, Vilnius University, Lithuania

The 34th ALTE meeting, hosted by the Department of Lithuanian Studies of Vilnius University in November 2007 was the first to be held in the Baltic States. As usual, the meeting consisted of two days of workshops, presentations and meetings for ALTE members and affiliates plus a conference day, this time held in the Parliament of the Republic of Lithuania, and open to the general public. The theme of the conference day was 'Language Testing in National Educational Systems: Communicating the Code of Practice to Stakeholder Groups' and was well attended by testers, teachers and other educationalists from the region as well as by ALTE members and others from all over Europe. The conference day was opened by Virginija Būdienė the Vice-minister of the Ministry of Education and Science and Vydas Gedvilas, a member of the Committee of Education, Science and Culture of the Parliament of the Republic of Lithuania.

Two presentations on the conference day were made by Cambridge ESOL staff. The first of these was made by Nick Saville (Director, Research and Validation Group) on *Comparing Quality: What Language Teachers Need to Ask Exam Providers*. Nick argued that teachers often have an important role in helping students to select a suitable exam and, to help them do so, awareness of the principles of assessment is of great benefit. He likened choosing a public exam to choosing a camera, where there are many choices and the key to making the right choice is understanding how different features may effect the uses the camera (or exam) can be put to. Nick then outlined the work which ALTE has done in providing tools with which stakeholders can assess exams, and focussed particularly on those used in relation to alignment to the CEFR and test quality. Dr Neil Jones (Principal Research Co-ordinator, Research and Validation Group) also made a presentation this time on formative assessment, entitled *Linking Learning and Assessment: A Can-Do Framework*. Neil spoke about the need to link the cognitive and social dimensions of validity. The cognitive dimension focuses on the process of learning and may be viewed as an input to learning; the social dimension relates to the output, or use of the learning and is often described using Can Do statements. Neil outlined some of the difficulties involved in such an undertaking, such as the institutional misuse of assessment, the conflicting roles of teachers and the poor professional development of teachers with respect to assessment. He argued that the solution was to develop a conceptual framework in which language learning (the input) and language use (the output) can be seen as different aspects of a single, continual process.

Also in the conference programme were a *Baltic Panel: Social Integration and Testing Titular Languages in Baltic States* and presentations on national language testing

policy in their respective countries were made by Gilles Breton, France; Cecilie Carlsen, Norway; Kate Green of the Department for Children, Schools and Families, UK – whose talk related to Asset Languages – and Gitte Østergaard Nielsen, Denmark. A round table followed this, with Kate Green, Henk Kuijper (the Netherlands), Nick Saville (UK) and Beate Zeidler (Germany) discussing issues around national testing policy and answering questions from the floor.

As mentioned above, the conference day was just a part of the whole ALTE meeting. Workshops were organised for the first two days of the meeting and were as follows:

- All Different – All Equal? Towards Cross-Language Benchmarking Using Samples of Oral Production in French, German and Italian
- Cross-Language Equating of Reading and Writing: an Experimental Workshop by Neil Jones (Cambridge ESOL)
- Relating Your Examinations to the CEFR: an Introduction to the Specification and Standardisation Phase of the Council of Europe Manual
- ALTE Auditor Training Workshop.

Apart from the final workshop, these workshops displayed a strong focus on aligning examinations and the CEFR. Neil's workshop employed a ranking task based on Thurstone's Paired Comparisons for exploring the links between written samples and tasks in different languages.

Details of forthcoming ALTE events can be found on the ALTE website: www.alte.org

Language Testing Forum 2007, Leicester, UK

The 2007 Language Testing Forum (LTF) was held in Leicester in November, with strong attendance from Cambridge ESOL staff. As every year, LTF hosted presentations on a wide range of language assessment topics. This year's guest lecture (which traditionally draws on expertise beyond the testing field) was given by Tim Jordan, Professor of Cognitive Neuroscience at the University of Leicester. Professor Jordan discussed the latest research findings on how the brain reads and what these findings tell us about the cognitive processes involved in reading.

Fumiyo Nakatsuhara from the University of Essex presented on conversational styles in group oral language tests. She explored the impact of a test-taker's own and his/her group members' characteristics (extroversion-introversion and proficiency levels) on conversational styles. The impact was also examined across three task types and two group sizes. The study could have implications for our understanding of the construct of group oral tests, the fairness in peer-peer tests and the development of rating scales.

Szilvia Papp from Cambridge ESOL gave a presentation on *Developing Can-Do statements for young learners in the under 14 age group* – a talk that was co-authored with Neil Jones (Cambridge ESOL). The talk presented Cambridge ESOL's research methodology in constructing a set of Can Do descriptors that are appropriate for assessing young candidates and that could be linked to the Common European Framework of Reference levels. Szilvia discussed

a number of related research issues, such as the pattern of developmental growth in English language proficiency in under 14-year olds, performance features that may be shared between young and adult learners across the CEFR levels and the cognitive demands posed by existing Can Do descriptors for young candidates. Ardeshir Geranpayeh, also from Cambridge ESOL, contributed a session that discussed some practical issues in using categorical data in structural equation modelling.

National Romanian Association of Teachers Conference, October 2007, Timisoara

Andy Chamberlain and Margaret Cooze attended the 8th National RATE Conference (Romanian Association of Teachers) in Timisoara, Romania from 26th–28th October. The theme of the conference was ‘Learning and Teaching English – The Neverending Story’. Margaret Cooze presented on the forthcoming updates to the FCE and CAE exams. In Romania more than 5000 candidates enter for CAE each year, many of them from state schools including those that teach through the medium of English.

The conference was attended by over 300 teachers and represented schools from all regions. The conference programme showcased a number of informative presentations and workshops on a diverse range of ELT fields. Codruta Goşa and Liana Gherdan gave a presentation on the success of the YLE exams in Western Romania highlighting the growth and popularity of the exams and the excellent preparation they provide for candidates who move upwards into the Main Suite exams. In addition Codruta Goşa and Luminiţa Frenţiu delivered an update on the extensive range of Cambridge ESOL examinations being taken in the region and in particular in the state school sector. One of the reasons for their success is the commitment and enthusiasm shown by the teachers.

Forthcoming conferences in Cambridge

Association of Language Testers in Europe 3rd International Conference, Cambridge, 10–12 April 2008

ALTE has received an unprecedented response to its call for papers for the ALTE 3rd International Conference. More than 200 papers from 189 institutions in 53 countries were proposed, and to reflect the multilingualism of the event, papers were accepted in German, English, Spanish, French and Italian. The conference will be held from 10–12 April 2008 at the University of Cambridge, hosted by University of Cambridge ESOL Examinations (Cambridge ESOL). The theme of the conference is ‘the social and educational impact of language assessment’ and will cover a variety of topics.

As well as an opportunity to learn and share, ALTE Cambridge 2008 will offer outstanding opportunities to meet and network with professionals from a wide range of associated fields from around the world. Among these opportunities will be the Conference Dinner on Friday 11 April to be held in the historic King’s College.

Some places may still be available and delegates can register online and book a wide selection of accommodation using the dedicated accommodation bureau with secure on-line reservation by credit card. A full programme can be found on the ALTE website: www.alte.org/2008

International Association for Educational Assessment Conference, Cambridge, 7–12 September 2008

As part of its 150th anniversary celebrations, Cambridge Assessment will host the 34th International Association for Educational Assessment (IAEA) annual conference in the UK from 7–12 September 2008. For further details please visit: www.iaea2008.cambridgeassessment.org.uk

Recent publications of interest

Studies in Language Testing

The last quarter of 2007 saw the publication of another title in the *Studies in Language Testing* series, published jointly by Cambridge ESOL and Cambridge University Press. Volume 25 by Anthony Green is entitled *IELTS Washback in Context: Preparation for academic writing in higher education* and is based upon the PhD dissertation which he completed at the University of Reading, UK, in 2003. The volume reports an empirical study to investigate the washback of the IELTS writing subtest on English for Academic Purposes (EAP) provision. The study examines dedicated IELTS preparation courses alongside broader programmes designed to develop the academic literacy

skills required for university study. Using a variety of data collection methods and analytical techniques, the research explores the complex relationship existing between teaching and learning processes and their outcomes. The role of IELTS in EAP provision is evaluated, particularly in relation to the length of time and amount of language support needed by learners to meet minimally acceptable standards for English-medium tertiary study.

Key features of the book include: a review of the literature on washback and on academic writing; exemplification of how innovative tools, such as neural network analysis, can be combined with more traditional statistical and qualitative techniques; insights into the learning and test-taking processes of learners of academic writing skills; and

recommendations for course provision and assessment design. This volume will be of direct interest to providers and users of general proficiency and EAP tests, as well as to academic researchers and graduate students interested in investigating test washback and impact. It will also be relevant to teachers, lecturers and researchers concerned with the development of EAP writing skills. More information is available at: www.cambridgeesol.org/what-we-do/research/silt.html

Publications by ESOL research staff

Issue 4/2 of the journal *Language Assessment Quarterly*, published in mid-2007, was a special issue devoted to differential item functioning in language assessment. Among other articles, it included a co-authored paper by Ardeshir Geranpayeh (Cambridge ESOL) and Anthony John Kunnan (California State University, Los Angeles) entitled *Differential Item Functioning in Terms of Age in the Certificate in Advanced English Examination*. When standardised English-language tests are administered to test takers worldwide, the test-taking population could be varied on a number of personal and educational characteristics such as age, gender, first language, and academic discipline. As test tasks and items may not always be prepared keeping this diversity of characteristics in mind, it is essential for test developers to continuously monitor their tests in terms of whether all test takers are receiving a fair test. This study examines whether the test items on the Listening paper of the Certificate in Advanced English (CAE) examination functioned differently for test takers from three different age groups. The main results showed that while statistical and content analyses procedures detected DIF in a few items, expert judges could not clearly identify the sources of DIF for the items.

Issue 4/4 of *Language Assessment Quarterly*, published in late 2007, contains an article co-authored by Thomas O'Neill, Chad Buckendahl, Barbara Plake and Lynda Taylor, entitled *Recommending a Nursing Specific Passing Standard for the IELTS Examination*. Professional registration and licensure testing programs (e.g. nursing) in the US and elsewhere face an increasing challenge of measuring the competency of internationally trained candidates, both in

relation to their clinical competence and their English language competence. To assist with the latter, professional licensing bodies often adopt well-established and widely available international English language proficiency measures, e.g. TOEFL or IELTS. In this context, the US National Council of State Boards of Nursing (NCSBN) sought to develop a nursing-specific passing standard on IELTS that all US jurisdictions could consider in their licensure decisions for internationally trained candidates. Findings from a standard setting exercise were considered by NCSBN's Examination Committee in conjunction with other relevant information to produce a legally defensible passing standard on the test. This article reports in detail on the standard setting exercise conducted as part of this policy-making process; it describes the techniques adopted, the procedures followed and the outcomes obtained. The study is contextualised within the current literature on standard setting. The article also describes the nature of the policy-making process to which the study contributed and discusses some of the implications of including a language literacy test as part of a licensure testing program.

Early 2008 sees the publication of the second edition of the *Encyclopedia of Language and Education*, edited by Nancy Hornberger and published by Springer. This new 10-volume edition is successor to the earlier 7-volume set, edited by David Corson and published by Kluwer in 1997. That first edition included a volume on *Language Testing and Assessment* edited by Caroline Clapham. The comparable volume in the latest edition (Volume 7) is edited jointly by Elana Shohamy and Nancy Hornberger and it contains 29 chapters covering the four topic areas of assessing language domains, methods of assessment, assessment in education, and assessment in society. Lynda Taylor and Fiona Barker contributed the chapter on *Using Corpora for Language Assessment*, a topic which was not covered at all in the 1997 edition of the Encyclopedia. Their chapter addresses early developments and major contributions in corpus-building and exploitation, work in progress and future directions in the field, as well as the challenges and difficulties it raises. They demonstrate how the application of corpora and corpus linguistics to the testing and assessment of L1 and L2 language proficiency is now well-established and has a promising future.

IELTS award news

We announce below the recipients of the 13th Round of the IELTS Joint-funded Research Program followed by a call for proposals for the next round, together with details of the winner of the 2007 IELTS Masters Award with a call for entries for the 2008 award.

IELTS Joint-funded Research Program

We are pleased to announce those successful research applicants who submitted proposals for Round 13 (2007/8) of the IELTS Joint-funded Research Programme. They are as follows: (see Table 1 overleaf)

Table 1: Studies funded under round 13 of the IELTS Joint-funded Research Programme

Researchers	Research Title
Roger Hawkey, Tony Green and Aylin Unaldi, The University of Bedfordshire, UK	An investigation of the process of writing IELTS academic reading test items
Gaynor Lloyd-Jones, Charles Neame and Simon Medaney, Cranfield University, UK	A multiple case study of the relationships between students' plural academic progress and IELTS scores at an international postgraduate university
Guoxing Yu, Pauline Rea-Dickins and Richard Kiely, The University of Bristol, UK	The cognitive processes of taking IELTS academic Writing Task 1
Andrea Dlaska, Ewan Dow and Sarah Michelotti, The University of Surrey, UK	From IELTS to graduation: An investigation into IELTS as a predictor of postgraduate academic success at the University of Surrey
Catherine Doherty, Margaret Kettle and Allan Luke, Queensland University of Technology, Australia	Examining academic spoken genres in university classrooms and their implications for the IELTS speaking test
Tim Moore, Janne Morton and Steve Price, The University of Melbourne, Australia	Construct validity in the IELTS academic reading test: a comparison of reading requirements in IELTS test items and in university study
Glenys Merrifield, GBM and Associates Australia	An impact study into the use of IELTS by professional associations and registration entities: The United Kingdom, Ireland and Canada
Michael Singh and Wayne Sawyer, The University of Western Sydney, Australia	The efficacy of IELTS in choosing potentially successful students for teacher education courses: What is an appropriate score?

Call for funding proposals: Round 14 of the IELTS Joint-funded Research Program

IDP: IELTS Australia and the British Council are once again making funding available for research projects in 2008/9. Each year an annual amount is set aside for external researchers to submit research proposals relating to the IELTS test. All IELTS research is managed by a Joint Research Committee which agrees on research priorities and oversees the tendering process. Researchers are now invited to submit funding proposals for Round 14, 2008/9. Details of the call for proposals for Round 14, together with guidance on topics and application forms, can be found on the IELTS website: www.ielts.org

IELTS Masters Award 2007 winner

The IELTS Research Committee, comprising the three IELTS partners: the University of Cambridge ESOL Examinations, the British Council and IDP: IELTS Australia met in November 2007 to review the shortlisted submissions for the IELTS Masters Award 2007. The winner was Talia Isaacs from McGill University in Montreal, Canada. Talia studied in the Department of Integrated Studies in Education and her supervisor was Dr Carolyn E Turner.

Talia's dissertation, entitled *Towards defining a valid assessment criterion of pronunciation proficiency in non-native English speaking graduate students* offers a valuable contribution to L2 pronunciation proficiency research. Talia Isaac's full abstract appears below:

This exploratory, mixed-design study investigates whether intelligibility is "enough", that is, a suitable goal and an adequate assessment criterion, for evaluating proficiency in the pronunciation of non-native English speaking graduate students in the academic domain. The study also seeks to identify those pronunciation features which are most crucial for intelligible speech.

Speech samples of 19 non-native English speaking graduate students in the Faculty of Education at McGill University were elicited using the *Test of Spoken English (TSE)*, a standardized test of spoken proficiency which is often used by institutions of higher learning to screen international teaching assistants (ITAs). Results of a fine-grained phonological analysis of the speech samples coupled with intelligibility ratings of 18 undergraduate science students suggest that intelligibility, though an adequate assessment criterion, is a necessary but not a sufficient condition for graduate students to instruct undergraduate courses as teaching assistants, and that there is a threshold level (i.e., minimum acceptable level) of intelligibility that needs to be identified more precisely. While insights about the features of pronunciation that are most critical for intelligibility are inconclusive, it is clear that intelligibility can be compromised for different reasons and is often the result of a combination of "problem areas" that interact together.

The study has some important implications for ITA training and assessment, for the design of graduate student pronunciation courses, and for future intelligibility research. It also presents a first step in validating theoretical intelligibility models which lack empirical backing (e.g. Morley 1994).

Talia will be presented with her award and a cheque for £1000 at the Language Testing Research Colloquium (LTRC) being held in Hangzhou, China from June 25–28, 2008.

Call for entries for IELTS Masters Award 2008

Each year the IELTS partners sponsor £1000 for the Masters level dissertation that makes the most significant contribution to the field of language testing. Submissions should be for dissertations written in partial or total fulfilment of the requirements for a Masters degree or its equivalent. Dissertations will only be considered eligible if they were submitted and *approved* by your university in 2007. The IELTS Research committee reviews the submissions and shortlists potential award winners.

Submission and evaluation procedures for 2008, along with details of the application process and timetable for submissions, can be found on the IELTS website: www.ielts.org