

# International assessment through the medium of English: analysing the language skills required

Stuart Shaw CIE

## Introduction

International assessments in a wide range of subjects are being prepared for and delivered through the medium of English. These are taken by many candidates whose first language is not English and increasingly by students who have participated in Content and Language Integrated Learning (CLIL) programmes in a range of different linguistic contexts. CLIL – defined as “an approach in which a foreign language is used as a tool in the learning of a non-language subject in which both language and the subject have a joint role” (Marsh in Coyle, 2006, p.1), involves “learning to use language and using language to learn” (Marsh and Lange, 2000). The CLIL approach consists of teaching a curricular subject through the medium of a language other than that which is normally used and operates in a range of contexts and is subject to varying interpretations. In CLIL programmes of learning, learners gain knowledge of the curricular subject while simultaneously learning and using the foreign language: curricular content leads language learning. Interaction in learning – a fundamental tenet of CLIL, is important because learners need to use and develop language *of* learning (the content); language *for* learning (peer interaction); and language *through* learning (for cognitive skills).

Typically students preparing for University of Cambridge International Examinations ('Cambridge') do so in very diverse linguistic and educational contexts, some following an entire curriculum in English, and others undertaking only one or two Cambridge examinations in parallel with qualifications from their own (non-English) national curriculum. The integration of two curricula in bilingual education programmes presents challenges for the schools and interesting issues for Cambridge. Cambridge is keen, therefore, to understand this context in order to evaluate the impact of this choice of education programme and particularly the role of assessment within it.

The Cambridge context raises a number of issues relating to *language awareness* (e.g. progression from basic interpersonal communication skills to cognitive academic language proficiency) and *assessment* (e.g. the level of English needed to access and succeed in international assessments).

The focus for the study described here is the International General Certificate of Secondary Education (IGCSE). The study adopts a two-phase methodology and involves an analysis of language use in Geography, History and Biology in order to (a) build a 'profile' of the language skills required and evidenced by IGCSEs and (b) determine whether any identifiable linguistic patterns adhere to different content, non-language IGCSEs.

## Language, educational and cognitive development

Students studying content subjects in a second language (L2) need to demonstrate competence not only in their familial linguistic background

(L1) but also within the educational community in which they are required to function or operate. This raises issues relating to bilingualism. Although it is not the purpose of this article to explore definitions of bilingualism, bilingualism is used here to refer to the use of two or more languages to operate in society, without regard to the level attained (see Mackey's [1968] use-based definition). Grosjean (1982, p.220) – with reference to the earlier work of Jespersen (1922), points out a child

*hardly learns either of the two languages as perfectly as he would have done if he had limited himself to one.*

Language acquisition has clear implications, therefore, for a learner's educational development:

*The brain effort required to master two languages instead of one certainly diminishes the child's power of learning other things which might and ought to be learnt.* (Jespersen 1922, in Grosjean, 1982, p.220)

## Language development

A number of linguistic idiosyncracies have been observed amongst students who exhibit language competence in two or more languages, particularly amongst bilinguals (Kelley, 1936; Tireman, 1955; MacLaughlin, 1978). These include limited vocabularies and grammatical structures, unusual word order, errors in morphology, hesitations and stuttering. MacLaughlin (1978) has argued that such difficulties are less to do with the process of bilingualism but more the fact that such children are forced to learn a second language in the school and do not have equal exposure to the two languages. An alternative thesis is offered by Skutnabb-Kangas and Toukomaa (1976) who have proposed the *Developmental Independence Hypothesis*. This hypothesis suggests that if the L1 is poorly developed, then focus on the L2 will impede the continued development of the L1. As a consequence, development of the L2 will be inhibited and lead to 'semi-lingualism'.

Cummins (1976) has suggested that children can – contingent upon teacher, home and community support, become bilingual at no cost to their L1. In his *Common Underlying Proficiency Theory* (Cummins, 1980), Cummins argues that the two languages used by an individual, though on the surface apparently quite separate and distinct, function through the same central cognitive system or as Baker asserts: “When a person owns two or more languages, there is one integrated source of thought” (Baker, 1996, p.147).

## Educational development

Some research indicates that learners who have been required to develop linguistic competence in two (or more) languages lack both interest and initiative and have, as a consequence, fallen behind educationally (Macnamara, 1966; MacLaughlin, 1978). According to MacLaughlin (1978), any educational impediment can be accounted for by testing

content in a second language over which the child has not yet developed sufficient command, combined with other minority ethnic factors such as low socio-economic status and negative attitudes of the majority group. In order to address the problem of insufficient command of L2, Cummins has proposed the Basic Interpersonal Communication Skills (BICS)/Cognitive Academic Language Proficiency (CALP) distinction.

The acronyms BICS and CALP are commonly used to discuss the language proficiency levels of students who are in the process of acquiring a new language. In an attempt to understand progression in students' learning of content and language, Cummins has shown how students need to progress from BICS (low cognitive demand, context embedded) towards CALP (high cognitive demand, context reduced). The distinction was intended to highlight the different time periods experienced by students to acquire conversational fluency in their L2 as compared to academic proficiency in that language. CALP is a language-related term which refers to formal academic learning, as opposed to BICS which are language skills needed in social situations. Typically, students develop proficiency in BICS well before they acquire a strong grasp of CALP: conversational fluency is often acquired to a functional level within about two years of initial exposure to the second language whereas development of academic aspects of the second language often takes between five and seven years (Cummins, 1981; Collier, 1987; Klesmer, 1994). As a consequence, students may give the appearance of being fully proficient and fluent, while still struggling with significant academic language deficiencies.

From a pedagogic perspective, the BICS/CALP distinction helps teachers support students to access cognitively challenging content material by embedding activities in a supportive context.

However, the BICS/CALP distinction is not without its detractors:

- the distinction reflects an autonomous perspective on language that ignores its position within social practices/power relations (Edelsky *et al.*, 1983; Wiley, 1996).
- CALP promotes a 'deficit theory' in that it attributes the academic failure of bilingual/minority students to low cognitive/academic proficiency as opposed to inappropriate education (Edelsky, 1990; Edelsky *et al.*, 1983; Martin-Jones and Romaine, 1986).
- CALP represents little more than 'test-wiseness' (Edelsky *et al.*, 1983).

The BICS/CALP distinction continues to engender debate. Notwithstanding the arguments, the distinction has had a longstanding effect on education and bilingual education in particular and is promulgated in strategic policy. For example, Tucker (1999 website) comments that the study carried out for the World Bank by Dutcher in 1994 concluded that:

*the best predictor of cognitive/academic language development in a second language is the level of development of cognitive/academic language proficiency in the first language and that cognitive/academic language skills, once developed, and content-subject material, once acquired, transfer readily from one language to another.*

Related to BICS/CALP are the concepts of *content-obligatory* and *content-compatible* language. When learning content through a second language, it is a requirement for a student to produce both content-obligatory and content-compatible language in a potentially wide range of subjects.

Content-obligatory language or specialist language is the language

that can be taught in the context of a particular subject and is essential to an understanding of content material. This is the subject-specific vocabulary, grammatical structures and functional expressions learners need in order to be able to learn about a curricular subject, communicate subject knowledge, and participate in interactive classroom tasks. In the context of History, for example, learners can discuss history either using general historical terms and phrases that are needed to operate within the subject but are not tied to a given period (e.g. collapse, defeat, democratic), or using words and phrases relating to the specific periods/events studied, which mainly amounts to nouns and proper nouns (e.g. conscription, hyperinflation, treaty).

Content-compatible language is language that can be taught naturally within the context of a particular subject matter and that students require additional practice with. This is non-subject-specific language which learners may have been exposed to and learned in their English language classes and which they can use in CLIL classrooms to communicate more substantively in the subject.

Examples of content-obligatory and content-compatible language in the context of Biology are shown in Appendix A.

### Cognitive development

The literature on the cognitive effects of language learning is mixed. Some research suggests that foreign language education increases cognitive development and positively influences academic achievement in other subjects. Stewart (2005) cites previous studies that found positive correlations between bilingualism and non-verbal measures of cognitive ability in young children. Grosjean (1982) notes that whilst some research indicates no effect on cognitive growth (Barik and Swain, 1976), other researchers have claimed negative effects (see Saer, 1926; Darcy, 1946). Lambert (1977) argues, however, that early IQ studies were beset with research methodology weaknesses (including not controlling for age, sex, socioeconomic background, educational opportunities, degree of bilingualism, matching on too few factors, lack of test adaptation for the linguistic minority).

Peal and Lambert in 1962 claimed French-English balanced bilinguals to be superior intellectually – scoring higher on both verbal and non-verbal IQ tests. However, the authors did concede that it is not clear whether intelligence is the reason for such an outcome. Others have also argued that bilinguals can have superior thinking abilities based on their dual linguistic systems. Garcia (2009) cites Vygotsky (1932) who contended that bilingual children had two ways to describe the world and so had more flexible interpretations. Garcia also notes work by Scott (1973) who reported more divergent thinkers amongst bilinguals when he told subjects to think of an object and say as many things as possible that they could do with it. Garcia notes that such studies show that bilingual children tend to give more responses, which are original and elaborate.

In attempting to resolve the conflict between the positive and negative effects, Cummins (1976) has suggested that there may be a threshold level of linguistic competence which a bilingual child must attain both in order to avoid cognitive deficits and allow the potentially beneficial aspects of developing bilingualism to influence their cognitive functioning. The 'Threshold' Theory was first put forward by Toukoma and Skutnabb-Kangas in 1977. It suggested that the development of two or more languages in a balanced bilingual person moves upward through three identifiable levels, crossing two distinct thresholds in between levels. According to this theory, positive cognitive advantages are only to be achieved when the first and second thresholds have been crossed.

# The International General Certificate of Secondary Education (IGCSE)

The focus of this study is the International General Certificate of Secondary Education (IGCSE).<sup>1</sup> The IGCSE is taken in a range of subjects at the end of a two-year course. At a similar and recognised level to the UK General Certificate of Secondary Education (GCSE), the IGCSE was developed for a global market, striving for non-UK centric contexts and awareness of second language needs. The IGCSE is open to schools from all over the world and is available twice a year in June and November. In many subjects there is an Extended and a Core Curriculum. The Extended Curriculum includes the material from the Core Curriculum, as well as additional, more advanced material. Each learner's performance is benchmarked using eight internationally recognised grades: Extended Curriculum: A\*, A, B, C, D, E; Core Curriculum: C, D, E, F, G.

## Research questions

This study sought to address the following questions:

- What level of English, according to the Common European Framework of References for Languages (CEFR), is needed to access and achieve in typical IGCSE assessments?
- What cognitive and academic language skills are needed to access and succeed at typical IGCSE assessments?

Key specific linguistic questions for both phases of the study were organised under three principal themes:

### *Lexical, structural and functional resources*

- What are the main language functions that students are being asked/demonstrating in their answers?
- Is there a pattern in the occurrence of structural forms of a particular type?
- Are examples of assessment specific vocabulary clearly comprehensible from syllabus guidelines?
- What are examples of subject specific vocabulary and what proportion of test questions mention or require responses involving subject specific vocabulary?
- Have candidates understood assessment specific vocabulary and effectively applied the requirements appropriately in their responses?

### *Expected and actual candidate performance*

- What writing skills required in mark schemes were anticipated/reflected in candidate responses?
- To what extent were candidates penalised by the ineffective use of subject specific vocabulary?
- What are the observations of the Principal Examiner on the use of language?
- How does candidate use of language compare with analysis of question papers and mark schemes?

### *Criteria task features relating to student performance*

- Is there evidence of undue cognitive reading demand made of candidates?
- What is the typical length, format and complexity of question input and rubrics/candidate responses?
- Is achievement linked to length of response?

## Methodology

The first phase of the study focused on Geography, History and Biology from the November 2008 and June 2009 sessions and entailed an analysis of syllabuses, question papers and mark schemes to allow a full overview of the qualification. In addition to analysing the June 2010 question papers and mark schemes, Phase 2 also involved an analysis of candidate performances, consultant and examiner reports.

Focus was on the written components (as opposed to practical or coursework components). In order to obtain varied perspectives on each IGCSE, four grade levels were sampled (A, C, E and F) from four linguistic backgrounds (Romance; Semitic; Sinitic; Slavic). As Biology includes a multiple choice paper and requires shorter written responses in candidate scripts, five candidates at each grade were studied, whereas in History and Geography three candidates were sampled at each grade.

The final data set comprised 74 Biology scripts; 47 History scripts; and 48 Geography scripts. Additional documentation was provided for the second phase of analysis in the form of reports on the issues of language written by senior examiners. These reports together with Principal Examiner insights enabled Phase 2 to be located in a broader context. (Principal Examiners are responsible for standards in the setting of question papers and the marking of examination scripts.)

## Findings

The findings are presented in terms of:

- the minimum level of English competence required to access and succeed at IGCSE;
- how the linguistic demands in the qualification might relate to the CEFR;
- the extent to which the language competence demonstrated could be defined as CALP.

In order to understand how the findings relate to the CEFR, a short description of the purpose and structure of the CEFR is provided.

Designed as a guideline to describe achievements of learners of foreign languages across Europe (and increasingly in other countries), the CEFR is a framework that provides a basis for the mutual recognition of language qualifications and enables awarding bodies to define and articulate language proficiency levels and interpret language qualifications.

The framework identifies six levels of potential language proficiency, two at basic language user level, namely A1 *Breakthrough* and A2 *Waystage*; two at independent user level: B1 *Threshold* and B2 *Vantage*, and two at proficient user level: C1 *Effective Operational Proficiency* and C2: *Mastery* level. The six reference levels are becoming widely accepted as the European standard for grading an individual's language proficiency. To illustrate these levels, CEFR global scale reference level descriptors

1. <http://www.cie.org.uk/qualifications/academic/middlesec/igcse/overview>

have been provided as Appendix B. The reference descriptors constitute a superordinate set of specifications, among nearly 60 scales provided by the CEFR (Council of Europe, 2001) to define different language skills, communicative purposes, contexts, activities, modes, etc.

The CEFR scales are intended to inform the development of language curricula, courses, tests and other forms of assessment, summative and formative, external and internal. The CEFR has growing relevance for language testers and examination boards, helping to define language proficiency levels and interpret language qualifications.

Specific findings from each of the two phases of the study are now reported.

## IGCSE History

The input language used in IGCSE History is of a high level. The language of the rubric falls mainly within the B2 level of the CEFR in terms of structural and lexical load. Although the rubrics and questions are generally expressed clearly using accessible language and could be understood by a B2 level student, the lexical input of the accompanying stimulus material is much higher and students would need to be at least C1 level to be able to process the text. There are many examples of structurally complex input including cleft sentences; organisation in terms of desired thematic prominence (rather than for accessibility or simplicity of structure); reported speech using a range of verb tenses, relative clauses and conditional structures.

Candidates need to be able to cope with a significant amount of subject-specific language, meaning that CALP is required. The question papers, and the source material which the questions refer to, contain a large amount of subject-specific vocabulary. Generally, this vocabulary falls into two lexical categories:

- **general historical terms and phrases** needed to operate within subject but are not tied to a given period (e.g. collapse, defeat, democratic and phrasal verbs such as set up, step in, take away)
- **specific lexis** that is linked to certain periods or topics (usually nouns and proper nouns such as conscription, colony, hyperinflation, dissidents, treaty)

Source texts may contain low frequency language and be challenging in terms of their 'authenticity'. Some of this material is complex (for example, the fronting of sentences with complex noun phrases) placing a high cognitive reading demand on candidates who are expected to quote from the material, and use it selectively in the exemplification of their argument. It is envisaged that candidates are prepared for this fact, and will also have in-depth knowledge of the historical period in question.

The use of cartoons and artwork in the input may pose challenges in terms of cultural non-familiarity though they may help to lessen the reading load. Their selection engenders interesting issues of accessibility and cultural relevance, and their appearance on papers may cause different challenges for candidates in different parts of the world.

Although candidates do not need to read source texts much longer than 250 words, they do, however, need to demonstrate a range of reading skills and strategies including careful reading at global level; careful reading at clause / sentence level; intensive reading of data; dealing with unfamiliar words and referencing skills (including exophoric referencing to link what they have read to a wider historical context).

The use of high-level input information used to set the scene for History questions suggests that emphasis is being placed on the top-down processing model of language or reading comprehension. This is a

model based on the belief that readers make sense of discourse by moving from the highest units of analysis to the lowest, and that comprehension is achieved by firstly activating background knowledge or schemata and setting the context.

Questions range from those requiring short answers (low tariff) to those requiring longer answers (high tariff), which ask for opinion, evaluation, justification and explanation with reference to source material. Short answers can be written at word or phrase/sentence level, but more open questions require longer, coherent answers usually consisting of more than one paragraph.

If the exemplification in the mark scheme for Paper 1 (consisting of questions selected from the 19th century and 20th century 'Core' topics) were to be seen as typical of the target output, candidates would be expected to produce language that is well above B2 level of the CEFR, even if the content-specific language is disregarded.

In terms of their written ability, IGCSE History candidates need to be able to demonstrate a range of writing skills. Students learning about History are required to be able to organise their ideas clearly, in order to present effective and balanced arguments that show evidence of evaluation and interpretation. They also need to demonstrate concision in certain questions and extended reasoning in others. They need to be able to quote judiciously from sources and exemplify claims from their knowledge. History teachers may need to teach this language or at least make learners aware of it in order for learners to be able to use it effectively

While accuracy of surface features such as spelling, word order, and grammar may not be fully mastered, candidates need to have a solid repertoire of structures, together with a wide vocabulary range. This will include many subject-specific terms and a number of nouns and proper nouns relating to the specific periods they have studied.

An important observation from the second phase is that low marks usually stem from deficiencies in the subject – lack of recall, inaccurate claims, unsupported assertions, one-sidedness, misinterpretation of question or sources, failure to evaluate, etc – rather than any obvious linguistic shortcomings.

## IGCSE Biology

The input language used in IGCSE Biology is not of a very high level. A student with B2 level English could do well on this qualification. There seems to be little or no requirement for detailed explanations or reasoned speculations, both of which would require an advanced level of English. Whilst students with B1 level English could access the paper and understand the questions they would not have a flexible enough command of English at their disposal to allow them to work at the speed required to complete the paper in the time given. Knowing the answer is the first step but having the language resource to describe processes/factors/differences with limited drafting time is a B2 level skill.

Generally, rubrics and questions are clearly written and a simple sentence structure is used (usually employing imperatives). The structures are often repeated. It is rare for any one part of a question to take up more than two lines and the layout is spacious and accessible.

Candidates have to read and understand a range of forms of input: graphical data (diagrams, tables), photographs, short/long questions, instructions (for the practical test). As with History, candidates need to employ a range of reading skills.

A number of different functional verbs are used for Biology, and each has a precise use and meaning though these subtle distinctions may not

be clearly understood by teachers and candidates who will have encountered these verbs previously in different language learning contexts. There is, therefore, an added level of challenge required to recognise the exact force of these verbs and produce what is required. Consequently, there is a far broader range of language functions involved in Biology than in, say, History and the subtleties underpinning the different verbs will have to be mastered, if candidates are to succeed.

The wording of questions is kept simple and structural forms are controlled within this. Gap-filling tasks are a good example of structural simplicity, and would be fully accessible to students from B2 level. Some question types involve processing and deduction. However, factual points are made clearly in single sentences which are then separated by a line space to assist candidates in their reading. This type of question is balanced by others with minimal text and which include the visual support of diagrams or illustrations.

On the multiple choice question Paper 1 (consisting of four-option multiple choice where candidates have 45 minutes to work through forty questions of differing formats, some including illustrations and others text and tabulated data), candidates clearly have to work at speed, reading efficiently. The reading load is not excessive on this paper though certain questions involve four full-sentence options.

The level of content specific vocabulary is very high across the papers. The learning of subject-specific terms for Biology is inextricably linked with the learning of the subject itself, in a way that is very different from, say, the learning of History.

Candidates often have the visual support of diagrams for a science subject (the level of graphical data input is high with a majority of the questions comprising graphical data in some form). However, there is inevitably a huge learning load, and all questions use subject-specific vocabulary, even if the responses do not always require it.

Language competence does not impact on Biology as much as in History. The Biology student is not required to produce long developed/reasoned answers: the mark scheme does not award marks for reasoning and development. Most of the answers requiring continuous prose are descriptions which can be done successfully with simple structures and key content-specific terminology.

Phase 2 reflects the findings of Phase 1: the language used in the Biology question papers is generally quite simple, with predictable structural forms and a limited range of command terms. However, the subject-specific vocabulary is much more demanding and key terms can be found in almost every question in each paper. Candidates who do not have a good grasp of this vocabulary would struggle to complete the questions.

Candidates are not required to produce long, detailed pieces of writing and many answers can consist of single words, short phrases or a few sentences at most. Where longer sentences are produced, most can be written using present simple or present continuous tenses, active and passive forms, basic conditional structures, comparatives or imperatives. Whilst candidates need to be able to produce these structures, conveying meaning appears to be more important than accuracy of expression.

Marks are only awarded for stating facts or identifying factors, reasons and so on – the style in which the answer is written does not matter. However, as well as naming things, stating facts or defining terms, candidates are required to interpret information and data, speculate, make suggestions and give detailed explanations, all of which are academic skills which need to be learnt in the classroom. Topics may not be immediately familiar to candidates and they may be required to make

connections to the subject matter they have learnt, draw conclusions or apply their scientific knowledge to a given situation. Therefore, CALP is important to some degree – candidates need to be able to assimilate information, know what type of information is required for each question, be able to make links, apply knowledge, and so on. Those candidates who give descriptions of what they see in a diagram rather than interpreting it would not be awarded marks; those who only state a fact but ignore the instruction to also give an explanation would be heavily penalised.

## IGCSE Geography

The input language used in IGCSE Geography is not of a very high level. Generally, a B2 level student would be able to cope with the vast majority of the rubrics, questions and input material. In the question papers, assessment-specific vocabulary appears in the rubrics and in the questions themselves, giving instructions and specifying the functional language which candidates are required to produce. Like History, candidates are required to identify from the rubrics the functional language required.

Candidates have to read and understand a range of forms which include graphical data (diagrams, bar charts, pie charts, maps, tables), photographs, short/long questions, short texts. The volume of graphical data is high but much of it can be understood only if the accompanying text is understood. In all three papers candidates are required to scan input material (whether it is a table, map or text extract) to locate answers. Candidates are also required to read intensively for detail. This entails reading a wide range of graphical data carefully; separating data from questions; reading numerical and other information from graphical data accurately; and moving between graphical data and text.

The papers contain a mixture of closed and open question types, requiring answers of varying length and format though the length of the answers is not specified. Overall, there is not a significant amount of extended text for candidates to read in any of the question papers, however all the papers consist of several questions, which each have a different number of sections and sub-sections. As a consequence candidates need to employ a variety of reading skills.

In Paper 1 (in which questions are resource based, involving problem solving and free response writing) they need to skim read the six questions in order to choose which three questions to answer. This involves reading the whole question with all its sections to check which information on which aspects of the topic is required for each section. Candidates must ensure that each section is answered and repetition/overlap of information is avoided. Candidates also need to read stimulus texts through before answering.

As papers do not have a standard format, candidates need to concentrate to read different question formats and different question types. Candidates may also need to deal with unfamiliar lexis which would entail deciding whether the unknown word is a key word and determining linguistic clues (using pictures/diagrams). Candidates need to be able to read the rubrics and questions carefully at clause and sentence level in order to be able to identify the type of response required (key words in the rubric) and what functional language to use in their answers. This can sometimes involve sophisticated recognition of textual patterns.

Like History and Biology, candidates are required to be able to cope with content-obligatory language with most questions containing some subject-specific vocabulary. Some of this vocabulary is not very high

frequency and may be challenging at this level. Whilst the level of content specific vocabulary is quite high – with some questions comprising higher-frequency language than others – all questions require candidates to understand subject-specific vocabulary and then to produce appropriate subject-related vocabulary in their answers.

There is an expectation that the candidate has flexible language resources to deal with a wide variety of question types. A B2 level student should be able to produce adequate responses, providing they have the lexical range, but CALP is required. Short answers require quite specific content-based language; longer answers need content knowledge but also a range of language to be able to describe, explain and draw conclusions, as well as the ability to write concisely.

Geography is a subject where students not only have to learn how to work with data, but also how to communicate in writing a wide range of concepts and ideas. Students with a good knowledge of Geography learnt in L1 would struggle to 'translate' this knowledge into English unless they had advanced language skills. Students who study Geography in the context of the English language would have a huge advantage when coming to these papers where language competence plays a key role. They would have learnt the subject while also learning to explain why things happen or might happen in English. C1 level students with a good knowledge of Geography and good data skills would perform well on this assessment. They would be able to write concisely for short answers, reformulate and develop ideas and speculate in longer answers, drawing on ideas learnt during the course as well as on evidence in the data on the paper. They would have language resources such that they could construct cohesive and coherent answers at the speed required (the mark scheme rewards development when longer answers are called for).

Analysis of scripts reveals that all candidates are able to attempt the majority of the questions. The main issues in terms of language use are (i) format of answers, that is, note form, bulleted lists, longer explanations; (ii) the range of language used; and (iii) the accuracy of the language used.

Candidates were able to use a good range of subject-specific vocabulary. Some of the vocabulary has a more general meaning but is relevant to and appropriate for the topics in the papers. In addition to using subject-specific vocabulary, candidates also demonstrated successful use of a range of general language structures and expressions.

Two issues of interest with regard to candidate performance are, first, the ability to produce developed answers and, linked to this, the ability to deal with questions requiring some form of speculation and judgement. To quote the comment from the Principal Examiner on the June 2010 Paper 4 (Alternative to Coursework which includes questions involving an appreciation of a range of techniques used in fieldwork studies):

*Weaker candidates scored on 'practical questions, such as drawing graphs' while candidates 'of higher ability' scored well on the 'more challenging sections requiring explanation and judgement, especially hypotheses'.*

## Discussion and conclusions

### IGCSE alignment to the CEFR

On the evidence of this study, candidates entering for the IGCSE History examination will be above B2 level and those attaining A and C grades will be at C1 or above.

Although many of the IGCSE Biology candidates are of a very high level and may even be bilingual, a minimum level of B2 on the CEFR is required. This is in part due to the high level of subject-specific language that they are required to cope with, but also because not all the topics are immediately familiar to candidates, the fact that some evaluation or synthesis of information is required, and that key points in explanations need to be made clearly and without ambiguity.

For IGCSE Geography, the level of output of grade A candidates is certainly C1 in terms of range, accuracy and control of collocation. Candidates scoring lower grades are writing at B2 level and sometimes below. Although the approach to accuracy is not explicit in the mark schemes, it is assumed that comprehensibility of the answer is crucial as there is evidence that answers with non-impeding errors and only very basic cohesion score marks for content. In terms of the questions requiring explanation, speculation and judgement, the level of language in successful answers is closer to C1 than B2. Explaining content in black and white terms can be done at B1/B2 level but to qualify ideas, to describe the colours in between – is an advanced language skill.

Therefore, it can be concluded that a minimum CEFR level of B2 is useful to access typical IGCSE subjects, and that a CEFR level of C1 could provide an added advantage of linguistic resources to be able to develop arguments needed for higher grades for Humanities subjects such as History and Geography. Each subject necessarily requires different types of CALP.

	IGCSE		
	History	Geography	Biology
<b>Overall CEFR alignment</b>	B2/C1	B2/C1	min B2
<b>CEFR User level</b>	Independent/ Proficient	Independent/ Proficient	Independent
<b>Requirement for CALP</b>	✓	✓	✓

### Supporting language claims underpinning the IGCSE

IGCSE claims an international reach and a local relevance as illustrated in the following quote taken from a current IGCSE Handbook:

*The syllabuses use international examples and avoid terminology only used in one country. Non-native speakers of English are always treated fairly. (Cambridge IGCSE, 2010, p.11)*

Fairness is concerned with "the consequences of testing for individuals, groups or society as a whole" (Davies *et al.* 1999, p.199) and is a social rather than a psychometric concept. Because fairness has no single meaning there is, therefore, no single definition. The *Standards for Educational and Psychological Testing* (AERA, APA, and NCME) note four possible meanings of fairness: (1) as requiring equal group outcomes; (2) as a lack of predictive bias; (3) as requiring that candidates have a comparable opportunity to learn the subject matter covered by the test; and (4) in terms of the equitable treatment of all candidates. One aspect of equitable treatment relates to the provision of reasonable accommodation for test takers with linguistic 'disadvantages'. To what extent then do the findings from this study substantiate the claim of fairness in terms of equitable treatment of all candidates?

The reading load in Biology is not high and it would seem that the quality of candidate responses depends less on time pressures than on

the ability to clearly express the required information. Thus marks are awarded according to the information stated rather than how it is expressed. Many of the questions across the Biology papers call for single words, short phrases or short descriptions. Therefore, candidates whose command of English is not fluent should still be able to complete the papers. The mark schemes do not describe specific writing skills that are required of candidates. However, it can be inferred from the mark schemes that single words and short phrases are acceptable, and that candidates do not gain extra marks by constructing complete sentences or longer, more coherent paragraphs. In general, there are few problems with candidates misunderstanding assessment-specific vocabulary – indeed many of them are using subject-specific vocabulary very effectively.

Whilst the language used in History is of a much higher level than Biology, examiners need to be congratulated for their tolerance of less than perfect English and their diligent processing of answers that are often dense, unparagraphed and written in challenging handwriting. The approach to marking appears to be positive rather than punitive, and any evidence, however thin, is likely to be sought out in order to raise an answer to the appropriate level of the mark scheme. Many candidates seem to perform effectively in English, which is a foreign language for them, apart from those few students who are fully bilingual. These candidates appear to be given every consideration both in terms of the questions they must answer and assessment of their responses.

Generally, there seem to be few problems with Geography candidates not understanding assessment specific vocabulary and most are able to provide answers appropriate to the question. Most candidates across the grades are able to complete all the questions and invariably with full answers. Even weaker candidates who score zero for many of their answers are able to write something for each question (sometimes at length and often with much irrelevance). In Paper 2 (based on testing the interpretation and analysis of geographical information and on the application of graphical and other techniques) and Paper 4 (the 'Alternative to Coursework' paper), many candidates, mostly those with a lower level of language range and accuracy, answered questions with phrases and bulleted lists, often with fractured grammatical structures. Stronger candidates produced full sentences and short paragraphs, almost always filling the lines provided for the response. Whilst the mark schemes make no reference as to whether both approaches are acceptable the assumption is that it is the content that counts, and not the style of the answer.

Interestingly, it is clear that candidates who use bulleted lists but have the linguistic resources to write full answers are penalising themselves unnecessarily. Those whose linguistic resources are not sufficient to support fuller answers can score satisfactorily on short-answer items (assuming subject knowledge) but cannot achieve maximum marks on questions requiring developed answers (and which often have higher totals of marks available).

In terms of relative time allowance, it is assumed that stronger candidates can produce longer and more cohesive text in the time given than weaker candidates. Sometimes, however, there is evidence of possible time advantage to candidates with knowledge of the correct answer and who opt for note form. In this case there is no evidence that providing lines for the answer guides candidates as to length; writing concisely is, however, a skill not always easy to acquire when writing in any language.

## Research informing practice

It is hoped that findings from this research will help to raise 'second language awareness' in all stages of development of question papers, mark schemes and examiner reports. Findings have already contributed to the question writing process: question setters need to be aware of potential language issues confronting an international candidature.

Outcomes will also inform the construction of a 'CALP guide' – *Language Awareness in Teaching: A Toolkit for Content and Language Teachers* (Chadwick, *in press*) – designed (a) for teachers of content subjects who teach to students for whom English is not their first language; (b) for English as a Second Language (E2L) teachers who teach students who take some of their content subjects in English in other departments of their school; and (c) for content teachers who teach students for whom English is their first language. (English may be the teacher's first or second language but in this case we can assume their proficiency in English.)

The function of the toolkit will be:

- to provide content teachers with a place to find the kind of language their students need support with when studying for their IGCSEs, and language that will enable their students to engage with the content subject more effectively. This language will be CALP that is useful for all academic subjects and examinations;
- to help content teachers become 'language aware';
- to include a rationale and strategies for supporting students with this language in the classroom;
- to provide guidance to E2L teachers on how they can support content teachers and students taking content subjects in English in their school;
- to provide E2L teachers with a resource that they can use to help plan and supplement their English lessons to be more effective across the curriculum.

## Future research

Building on the research reported here future studies will attempt to assess the impact of linguistic complexity and language accessibility on candidates taking international A level examinations designed for 16–18 year olds. The research is designed to comprise three phases. In phase 1, the marks obtained by each student for each sub-question on the exam papers for a random sample of at least 200 scripts for A level Geography and A level Physics will be collected and keyed into data spreadsheets. The data sets will be used to conduct a number of statistical analyses to describe question functioning for both whole questions and question parts using traditional and item response analyses. In phase 2, questions that statistical analyses suggest are performing in 'unexpected' ways (extremes of difficulty; reverse thresholds, a number of overfitting and underfitting items) will be explored further using textual and discourse analytic techniques in order to determine whether the questions present problems for international candidates and, more importantly, why these questions might be problematic. In the final phase of the research, students studying in their second year of A level Geography and A level Physics from a range of linguistic backgrounds will be asked to engage with the input language of questions identified in phase 1 and to comment on their linguistic complexity. Triangulation of textual analysis and think-aloud protocols will provide a powerful means to explore

complex syntactic and lexical features that challenge English language learners. Through the 'voices' of students, this work will scrutinise the appropriateness of inferences about English language learners' content knowledge based on linguistically complex test items.

More research is needed into ways of making academic content more accessible and meaningful to students in bilingual programmes, particularly in areas/subjects considered to be challenging when learning academic content occurs through the second language.

The research findings in respect of 'transfer' tend to support the positive rather than the negative: although more research is needed, the literature points to some evidence for transfer of skills across languages (academic skills, subject knowledge skills, literacy skills).

There is also an urgent need to develop effective bilingual assessment methods that reflect classroom practices of using two (or more) languages for teaching and learning – methods that move away from the notion of monolingual assessment and testing bilinguals as if they were two monolinguals – so that bilingual children are given the opportunity to show their proficiency and competences in both languages.

## References

- Baker, O. (1996). *Foundations of Bilingual Education and Bilingualism*. 2nd Edition. Clevedon: Multilingual Matters.
- Barik, H. & Swain, M. (1976). A longitudinal study of bilingual and cognitive development. *International Journal of Psychology*, **11**, 4, 251–263.
- Cambridge IGCSE. (2010). The Cambridge IGCSE Handbook. University of Cambridge International Examinations <http://www.cie.org.uk/search?searchfield=handbook%20igcse>.
- Collier, V. P. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly*, **21**, 617–641.
- Council of Europe. (2001). *Common European Framework of Reference for Languages*. Cambridge: Cambridge University Press.
- Coyle, D. (2006). Developing CLIL: Towards a Theory of Practice. Monograph 6, APAC Barcelona.
- Coyle, D. (2007). Content and Language Integrated Learning: Towards a Connected Research Agenda for CLIL Pedagogies. *International Journal of Bilingual Education and Bilingualism*, **10**, 5.
- Coyle, D. (1999). Theory and planning for effective classrooms: supporting students in content and language integrated learning contexts. In: J. Masih (Ed.) *Learning through a Foreign Language*. London: CILT.
- Cummins, J. (1976). The Influence of Bilingualism on Cognitive Growth. Reproduced in Baker, C. and Hornberger, N. H. (2001) from *Working papers on Bilingualism*, April 1976, 1–43, Ontario Institute for Studies in Education.
- Cummins, J. (1980). The Entry and Exit Fallacy in Bilingual Education. Reproduced in Baker, C. and Hornberger, N. H. (2001) from *NABE Journal*, **4**, 25–60.
- Cummins, J. (1981). Age on arrival and immigrant second language learning in Canada. A reassessment. *Applied Linguistics*, **2**, 132–149.
- Cummins, J. (1982). Tests, Achievement and Bilingual Students. Reproduced in Baker, C. and Hornberger, N.H. (2001) from *FOCUS*, February 1982, **9**, 1–7, National Clearinghouse for Bilingual Education, Washington DC, USA.
- Cummins, J. (1986). Empowering minority students: A framework for intervention. *Harvard Educational Review*, **56**, 18–36.
- Cummins, J. (1996). *Negotiating identities: Education for empowerment in a diverse society*. Los Angeles: California Association for Bilingual Education.
- Cummins, J. & Swain, M. (1983). Analysis-by rhetoric: reading the text or the reader's own projections? A reply to Edelsky et al. *Applied Linguistics*, **4**, 22–41.
- Darcy, N. (1946). The effect of bilingualism upon the measurement of the intelligence of children of preschool age. *Journal of Educational Psychology*, **37**, 21–44.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. & McNamara, T. (1999). *Dictionary of Language Testing*, Studies in Language Testing 7. Cambridge: UCLES and Cambridge University Press.
- Dutcher, N. & Tucker, G.R. (1994). *The Use of First and Second Languages in Education: A Review of Educational Experience*. Washington DC, USA: World Bank.
- Edelsky, C. (1990). *With literacy and justice for all: Rethinking the social in language and education*. London: The Falmer Press.
- Edelsky, C., Hudelson, S., Altwerger, B., Flores, B., Barkin, F., Jilbert, K. (1983). Semilingualism and language deficit. *Applied Linguistics*, **4**, 1, 1–22.
- Garcia, O. (2009). *Bilingual Education in the 21st Century: A global perspective*. Oxford: Wiley-Blackwell.
- Grosjean, F. (1982). *Life with Two Languages*. Cambridge, Mass.: Harvard University Press.
- Jespersen, O. (1922). *Language, its Nature, Development and Origin*. London: George Allen and Unwin Ltd.
- Kelly, V. (1936). Reading abilities of Spanish and English speaking pupils. *Journal of Educational Research*, **29**, 209–211.
- Klesmer, H. (1994). Assessment and teacher perceptions of ESL student achievement. *English Quarterly*, **26**, 3, 5–7.
- Lambert, W. E. (1977). The effects of bilingualism on the individual: Cognitive and sociocultural consequences. In: P.A. Hornby (Ed.), *Bilingualism: Psychological, social, and educational implications*. 15–27. New York: Academic Press.
- McLaughlin, B. (1978). *Second Language Acquisition in Childhood*. Hillsdale, New Jersey: Erlbaum.
- Macnamara, J. (1966). *Bilingualism and Primary Education*. Edinburgh: Edinburgh University Press.
- Marsh, D. & Lange, G. (2000). *Using Languages to Learn and Learning to use Languages*. Jyväskylä, University of Jyväskylä, Finland: UniCOM.
- Martin-Jones, M. & Romaine, S. (1986). Semilingualism: A half-baked theory of communicative competence. *Applied Linguistics*, **7**, 1, 26–38.
- Peal, E. & Lambert, W.E. (1962). The Relation of Bilingualism to Intelligence. *Psychological Monographs* **76**, 27, 1–23.
- Saer, D. J. (1923). The effects of bilingualism on intelligence. *British Journal of Psychology*, **14**.
- Scott, S. (1973). *The Relation of Divergent Thinking to Bilingualism: Cause or Effect?* Montreal: McGill University.
- Stewart, J. (2005). Foreign Language Study in Elementary Schools: Benefits and implications for achievement in Reading and Math. *Early Childhood Education Journal*, **33**, 1, 11–16.
- Skutnabb-Kangas, T. (1977). Language in the process of cultural assimilation and structural incorporation of linguistic minorities. In: C. C. Elert et al. (Eds.), *Dialectology and Sociolinguistics*. 191–199. UMEA: UMEA Studies in the Humanities.
- Skutnabb-Kangas, T. & Toukomaa, P. (1976). *Teaching migrant children's mother tongue and learning the language of the host country in the context of the sociocultural situation of the migrant family*. Helsinki: The Finnish National Commission for UNESCO.
- Tireman, L. (1955). Bilingual child and his reading vocabulary. *Elementary English*, **32**, 33–35.
- Tucker, G.R. (1999). *A Global Perspective on Bilingualism and Bilingual Education*, Carnegie Mellon University, <http://www.cal.org/resources/Digest/digestglobal.html>
- Vygotsky, L. S. (1932). Predislovie [Preface] In: Piaget J. Rech i m'shieme rebenka [*The speech and thinking of the child*]. 3–54. Leningrad: Uchpedgt.
- Wiley, T.G. (1996). *Literacy and language diversity in the United States*. Washington, DC: Center for Applied Linguistics and Delta Systems.



## Appendix A: Comparison of content-obligatory and content-compatible Biology language

<i>Content-obligatory language</i>	<i>Content-compatible language</i>
<ul style="list-style-type: none"> <li>● to describe leaves: 'waxy'; 'spikes'; 'cuticle'</li> <li>● to describe environmental problems: 'deforestation'; 'global warming'; '(bio)degradable'; 'the ozone layer'; 'endangered'; 'fossil fuels'; 'earthquakes'; 'drought'</li> <li>● to describe laboratory experiments: 'test tubes'; 'goggles'; 'pestle and mortar'; 'precipitate'; 'ethanol'; 'iodine'; 'Benedict's solution'; 'control'</li> <li>● to discuss use of fertilisers: 'eutrophication'</li> <li>● to explain the blood system: 'valves'; 'backflow'; '(oxy)haemoglobin'; 'deoxygenated'</li> <li>● to explain plant growth: 'germinate'; 'to wilt'</li> <li>● to describe teeth: 'molars'; 'incisors'; 'canines'; 'cusps'; 'dentine'; 'enamel'; 'root'</li> <li>● to identify parts of the human eye: 'cornea'; 'iris'; 'lens'; 'suspensory ligament'; 'yellow spot/fovea'; 'blind spot'</li> </ul>	<ul style="list-style-type: none"> <li>● adjectives or verbs with dependent prepositions: e.g. 'resistant to'; 'suffer from'; 'give off (energy)'; 'react to'; 'respond to'; 'immune to'; 'exposed to'; 'dependent on'; 'protect from'; 'fight off (disease)'; 'adapt to'; 'cut down (trees)'; 'consist of'</li> <li>● phrasal verbs: e.g. 'to break down (a substance)'; 'to carry out (a test)'; 'to set up (an experiment)'; 'to speed up' (photosynthesis/a reaction)</li> <li>● verb-adverb collocations: e.g. 'increased exponentially'; 'rises dramatically'</li> <li>● verb-noun and adjective-noun collocations: e.g. 'to have an adverse effect on...'; 'weaken their immunity'</li> </ul>

## Appendix B: Common Reference levels - Global Scale

<b>Proficient User</b>	C2	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
<b>Independent User</b>	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.
<b>Basic User</b>	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

Council of Europe, 2001, p.24