

Державний вищий навчальний заклад
“Українська академія банківської справи
Національного банку України”
Кафедра іноземних мов

ОЦІНЮВАННЯ ВМІНЬ ПИСЬМА

ASSESSING WRITING

Навчально-методичний
посібник

Для викладачів
кафедри іноземних мов

Суми
ДВНЗ “УАБС НБУ”
2013

УДК 811.111(073)
О-93

Рекомендовано до видання науково-методичною радою Державного вищого навчального закладу “Українська академія банківської справи Національного банку України”, протокол № 4 від 25.12.2012.

Розглянуто та схвалено на засіданні кафедри іноземних мов, протокол № 4 від 11.12.2012.

Укладач

Л. В. Гнаповська, кандидат філологічних наук, доцент,
ДВНЗ “Українська академія банківської справи НБУ”

Рецензенти:

О. В. Ємельянова, кандидат філологічних наук, доцент,
Сумський державний університет;

І. А. Бокун, кандидат філологічних наук, доцент,
ДВНЗ “Українська академія банківської справи НБУ”

Assessing Writing = Оцінювання вмінь письма [Текст] : навчально-методичний посібник / Державний вищий навчальний заклад “Українська академія банківської справи Національного банку України” ; [уклад. Л. В. Гнаповська]. – Суми : ДВНЗ “УАБС НБУ”, 2013. – 23 с.

Навчально-методичний посібник розроблений відповідно до навчальних програм курсів “Іноземна мова”, “Ділова іноземна мова” та “Ділова іноземна мова за професійним спрямуванням”. Запропоновані в посібнику матеріали стануть у нагоді викладачам під час розробки завдань для оцінювання вмінь письмової англomовної комунікації. Подані в посібнику рекомендації стосовно використання шкали оцінювання письма як продуктивного виду мовленнєвої діяльності покликані озброїти викладачів знаннями про валідні інструменти вимірювання рівня сформованості писемної комунікативної компетенції студентів.

Призначений для викладачів кафедри іноземних мов.

УДК 811.111(073)

© ДВНЗ “Українська академія банківської справи
Національного банку України”, 2013

CONTENTS

INTRODUCTION	4
PART 1. DESIGNING TASKS FOR ASSESSING SKILLS OF WRITING	5
1.1 The rubric	5
1.2 The prompt	6
PART 2. DESIGNING SCALES FOR ASSESSING WRITING ABILITY	8
2.1 Holistic scales.....	8
2.2 Analytic scales.....	10
PART 3. INTERPRETING ASSESSMENT CRITERIA	14
3.1 Task Achievement (TA).....	14
3.2 Coherence and Cohesion.....	15
3.3 Grammar.....	18
3.4 Vocabulary	20
REFERENCES	23

INTRODUCTION

The aim of modern English examinations is to give language learners tasks that require them to use language as similarly to real-life situations as possible. Such tasks present candidates with a variety of lifelike writing situations that allow them to perform different language functions to achieve their communicative goals. The tasks therefore present carefully designed contexts in which test-takers have to address topics relevant to them to achieve specific communicative purposes.

The writing tasks administered in modern English examinations attempt to ensure that candidates produce their own texts rather than reproduce memorized ones. The aim is to elicit language that is suitable for getting the student's message across. In modern English examinations more emphasis is laid on successful communication than on accuracy: a candidate does not have to produce a text that is perfectly accurate in order to be understood and communicative.

The handbook is intended to help teachers who have to administer and design tests of writing that meet the standards of modern European English examinations. It gives an overview of the composing process in order to describe the ability that a test of writing measures, provides an introduction to designing writing tasks, describes the main types of rating scales, and systematically explores the criteria most frequently used in the assessment of writing skills.

PART 1

DESIGNING TASKS FOR ASSESSING SKILLS OF WRITING

In order to be able to evaluate a writing assessment task, we need to know two basic things:

- the criteria that such a task must meet;
- how we can check whether it is a functioning task or not.

The two basic parts that can be distinguished in an assessment task are the *rubric* and the *prompt*. Although they have separate functions, in some tasks the rubric and the prompt are combined into one coherent mix of instructions and stimuli for the students, so their elements are intermingled.

1.1 The rubric

The rubric can contain information on:

- what exactly the task will assess (e.g. *This is a test of your ability to write a coherent and grammatically correct argumentative composition or Your letter will be evaluated on the basis of content, expression, register, structure and mechanics*);
- how students should go about completing the writing task (e.g. *Write the answer on the answer sheet*);
- the time given for the completion of the task;
- the word limit set for the task;
- how the task is marked (e.g. *Task 1 carries more weight in marking than Task 2*).

The function of the rubric is therefore to give procedural information.

The two pieces of information that must be given in a rubric are the time allotted for the completion of the task and the amount of text, usually expressed in number of words that test takers are required to produce. Test takers need to know how much time is given to do the task and how many words they have to write if they are to plan, write and hopefully revise their script successfully.

A rubric must provide all the information that is necessary for the students to understand what they have to do. However, this is not to be interpreted as meaning the longer the rubric the better. A task intended for the assessment of writing ability must not turn into a major reading activity; therefore, the rubric should be as short as clarity allows. Furthermore, the rubric must be clearly worded to make sure that it is understandable. The language used may be student's mother tongue, but in modern European examinations the rubric is usually written in the same language as the student is tested on. Therefore, the rubric must be written in simple language and the information presented in it must be logically organized and ordered.

1.2 The prompt

The exact nature and content of the script students are expected to compose are specified in the prompt. The prompt is the stimulus in response to which students write the script, so it must be understandable, rich, motivating and encouraging. The function of the prompt is to provide information about the context, content, and purpose of writing.

There are three basic prompt formats.

The **base prompt** is short and formulates the whole task in one or two sentences. Most one-sentence composition prompts are base prompts, e.g.

Do you agree or disagree with the opinion that men make better managers than women? Why?

Such prompt sets the topic and may specify the main rhetorical mode of the expected script; however, it does not provide information about the writing context and the exact content to be developed. It most resembles writing tasks set by language teachers who want to check whether their students can apply their knowledge about a particular grammar point in practice or whether they have learnt certain vocabulary items.

The **framed prompt**, on the contrary, is longer and helps students to interpret the task by placing it into a richer context. It usually describes a situation from which candidates can understand who they are, why they are writing, what they are writing about, and who will read their texts, e.g.

Your company has decided to look into using a local venue for holding in-house training for company staff. Your boss has asked you to research and then recommend a suitable venue, for HR Director to approve.

*Write your **proposal**, including the following:*

- *a summary of the types of training your company runs*
- *your criteria for selecting venue*
- *an evaluation of two or more local venues*
- *a recommendation for a local venue*

The contexts described in such prompts present contexts familiar to test-takers or the ones they can imagine themselves in. they are therefore modelled on real-life situations and are much preferred by task designers since life is an exhaustible source of authentic situations in which writing occurs.

The last type of prompt is the **text-based prompt** and it may include one or more texts of varying length to which students respond (e.g. a letter, an advertisement, brochures), ideas which they are expected to incorporate into their scripts (e.g. notes prepared for an article, or list of ideas jotted down to be included in a letter), or visual information presented in the form of images (e.g. a graph presenting certain trends). In some writing books the tasks that require the writer to compose a text in response to such stimuli are called *transactional writing tasks*. Text-based prompts are frequently used to test writing skills because they allow the creation of a large variety of stimulating lifelike writing tasks, e.g.

You see the job advertisement below in a national newspaper and decide to apply for it. Write a covering letter to send with your CV.

MARKETING MANAGER

JTC is one of the country's most established distribution companies. We are currently looking for a highly-motivated and experienced Marketing Manager to join our organization. The successful candidate will lead a team of more than 50 staff throughout the region. We expect you to have a professional qualification and experience of brand management, market analysis, advertising and communications. You will also be creative and willing to take responsibility for extending our current market

PART 2

DESIGNING SCALES FOR ASSESSING WRITING ABILITY

A rating scale is a means of defining the standards that should apply when a rater judges a learner's performance. A scale is a measuring instrument which defines what it is that is being measured. A scale can be seen as a line, ranging from a very low or weak performance to a very high (excellent) performance. Scales are divided into a number of points. These points may be labeled as numbers (e.g. 1 to 8), or with adjectives like *excellent*, *very good*, *weak*, *very poor* etc.

However, the problem with such labels is that they will mean different things for different raters. Therefore, in order to ensure fair marking, the points on the scale are usually described in words, in order to guide the marker in deciding which level to award the script.

Scales that do not have such descriptors are called **impression scales**, and the marking of scripts using such scales is called **general impression marking**. Such marking is common in traditional examinations, often using a scale from 1 to 5, where 1 is Fail and 5 is High Pass. However, such marking is usually very unreliable – different raters will assign different marks to the same script, depending upon how they have interpreted the label of the level on the scale.

Modern European exams do not use such impression scales. Rather they use a rating scale (often called a scoring rubric) which defines the criteria to be used when awarding a level to the script. There are two main types rating scales: holistic (also referred to as global) and analytic scales.

2.1 Holistic scales

Holistic scoring means assigning of a single score which is based on the overall impression of the script, guided by the descriptions of each level. Candidates are placed at a certain level on a scale. Writers of holistic scales provide overall descriptions of writing ability and include different features in a level, or band, at the same time. Such features may include reference to content, organization, grammar, vocabulary and mechanics. Raters are free to decide which feature will have the biggest influence on their decision to award a score to a script. Some may feel that in one case, grammar is more important than organization, whereas another rater may decide the reverse.

The essays are scored on a 6-band holistic scale in the TOEFL Test of Written English (TWE), shown as Figure 7.1 below.

Figure 1 – TOEFL Writing Scoring Guide

- 6** An essay at this level
- effectively addresses the writing task
 - is well organized and well developed
 - uses clearly appropriate details to support a thesis or illustrate ideas
 - displays consistent facility in the use of language
 - demonstrates syntactic variety and appropriate word choice
- 5** An essay at this level
- may address some parts of the task more effectively than others
 - is generally well organized and developed
 - uses details to support a thesis or illustrate ideas
 - displays facility in the use of language
 - demonstrates some syntactic variety and range of vocabulary
- 4** An essay at this level
- addresses the writing topic adequately but may slight parts of the task
 - is adequately organized and developed
 - uses some details to support a thesis or illustrate an idea
 - demonstrates adequate but possibly inconsistent facility with syntax and usage
 - may contain some errors that occasionally obscure meaning
- 3** An essay at this level may reveal one or more of the following weaknesses:
- inadequate organization and development
 - inappropriate or insufficient details to support or illustrate generalizations
 - a noticeably inappropriate choice of words or word forms
 - an accumulation of errors in sentence structure and/or usage
- 2** An essay at this level is seriously flawed by one or more of the following weaknesses:
- serious disorganization or underdevelopment
 - little or no detail, or irrelevant specifics
 - serious and frequent errors in sentence structure or usage
 - serious problems with focus
- 1** An essay at this level
- may be incoherent
 - may be undeveloped
 - may contain severe and persistent writing errors
- 0** An essay may be rated 0 if it
- contains no response
 - merely copies the topic
 - is off-topic, is written in a foreign language, or consists only of keystroke characters

2.2 Analytic scales

The other type of scales used in international examinations are analytic. Analytic scales rate scripts on several criteria separately, e.g. content, organization, language use (grammar), vocabulary, mechanics (spelling, punctuation) etc. the script is given a mark for each separate criterion and the final score awarded to the script – the final grade – is a composite of the assessments in respect of each criterion. This type of scales is particularly useful for diagnosing students' strengths and weaknesses in different areas.

If we compare the two types of scales, there are advantages and disadvantages to each. Holistic scoring is faster than analytic scoring and reflects an authentic reader's personal reaction to a text since readers often make judgements about texts based on an overall impression. However, experts warn of the dangers of holistic scoring, namely that the rater's judgement might be affected by just one or two aspects of the script, and that this may vary from rater to rater, thereby affecting the inter-rater agreement.

Analytic scoring is probably better for assessing foreign language scripts in particular, as language learners may show an uneven profile across different aspects of writing. For example, a script may have excellent content with bad grammar, or good grammar with weak organization. Analytic scoring takes longer than holistic scoring, but it is usually much more reliable, provided, of course, that raters have been trained to use the scales. Consequently, most international examinations use some form of analytic rating scale.

Not all the information necessary for rating can be included in a rating scale. Such scales have to be used together with **guidelines for raters** and there are often task-specific **descriptions of content points and requirements**, which will vary from task to task.

Shown below is the analytic writing scale developed for the purposes of assessing writing skills while teaching English/Business English at Ukrainian Academy of Banking of the National Bank of Ukraine (see Fig. 2). We also provide the Guidance for Raters (see Fig. 3) that accompanies the scale.

In what follows, we describe briefly how the rating scale and the guidance are expected to be applied in the UAB context.

Figure 2 – The UAB Analytic Writing Rating Scale

SCORE	CRITERIA			
	TASK ACHIEVEMENT	COHERENCE AND COHESION	GRAMMAR	VOCABULARY
5	<ul style="list-style-type: none"> • All content points fully elaborated • Meets text type requirements completely 	<ul style="list-style-type: none"> • Fully coherent text • Text cohesive on both sentence and paragraph level 	<ul style="list-style-type: none"> • Wide range of structures • Occasional inaccuracies that do not hinder/disrupt communication 	<ul style="list-style-type: none"> • Wide range of general and professional vocabulary • Accurate vocabulary communicating clear ideas • Fully relevant to content
4	<ul style="list-style-type: none"> • Most content points elaborated • All content points mentioned • Occasional inconsistencies in text type requirements 	<ul style="list-style-type: none"> • Good sentence-level cohesion • Text mostly coherent and cohesive on paragraph level 	<ul style="list-style-type: none"> • Good range of structures • Occasional inaccuracies that hinder/disrupt communication 	<ul style="list-style-type: none"> • Good range of general and professional vocabulary • Occasionally inaccurate vocabulary communicating mainly clear ideas • Overall relevant to content
3	<ul style="list-style-type: none"> • Many content points elaborated • Most content points mentioned • Some inconsistencies in text type requirements 	<ul style="list-style-type: none"> • Text cohesive enough on sentence level • Occasional lack of paragraph-level coherence and cohesion 	<ul style="list-style-type: none"> • Adequate variety of structures • Some inaccuracies that hinder/disrupt communication 	<ul style="list-style-type: none"> • Fair range of vocabulary • Frequently inaccurate vocabulary communicating some clear ideas • Occasionally irrelevant to content
2	<ul style="list-style-type: none"> • Some content points elaborated • Many content points mentioned • Many inconsistencies in text type requirements 	<ul style="list-style-type: none"> • Some sentence-level cohesion • Frequent lack of paragraph-level coherence and cohesion 	<ul style="list-style-type: none"> • Limited range of structures • Frequent inaccuracies that hinder/disrupt communication 	<ul style="list-style-type: none"> • Limited range of vocabulary • Frequently inaccurate vocabulary communicating few clear ideas • Occasionally relevant to content with some chunks lifted from prompt

Continued Figure 2

SCORE	CRITERIA			
	TASK ACHIEVEMENT	COHERENCE AND COHESION	GRAMMAR	VOCABULARY
1	<ul style="list-style-type: none"> No content point elaborated Some content points mentioned Does not meet text type requirements 	<ul style="list-style-type: none"> Text not coherent Lack of sentence- and paragraph-level cohesion 	<ul style="list-style-type: none"> No range of structures Mostly inaccurate 	<ul style="list-style-type: none"> No range of vocabulary Mostly inaccurate vocabulary communicating ideas that are not clear enough Mostly irrelevant to content with several chunks lifted from prompt
0	No assessable language	No assessable language	No assessable language	No assessable language

Figure 3 – The UAB Analytic Writing Rating Scale: Guidelines for Raters

CRITERIA FOR ASSESSMENT	CHECK	MAKE SURE
TASK ACHIEVEMENT	<p><i>Depth of coverage</i></p> <ul style="list-style-type: none"> Which content points are elaborated? Which content points are mentioned? <p><i>Text type requirements – task specific</i></p> <ul style="list-style-type: none"> Are the text-specific conventions observed? 	<ul style="list-style-type: none"> Content points elaborated with the most detail, not just mentioned briefly Thoughts and ideas are relevant and original. There are no irrelevant parts that do not belong in the text Stylistically appropriate (formal / informal) language is used Layout conventions of the text type are observed
COHERENCE AND COHESION	<p><i>Organization and linking of ideas</i></p> <ul style="list-style-type: none"> Is the script coherent? Is the script cohesive? 	<ul style="list-style-type: none"> Ideas are clearly organized and follow one another logically The relationship between sentences and their parts are marked clearly and correctly The linking devices used are varied and appropriate

Continued Figure 3

<p>COHERENCE AND COHESION</p>	<p><i>Paragraphing</i></p> <ul style="list-style-type: none"> • Does the script need to be and is it divided into paragraphs? <p><i>Punctuation</i></p>	<ul style="list-style-type: none"> • Ideas are organized in the way that one subtopic is developed into one paragraph • Paragraphs are properly indicated: they are either block or indented • The relationship between paragraphs are marked clearly and correctly • Appropriate punctuation marks are used correctly
<p>GRAMMAR</p>	<p><i>Grammatical range</i></p> <ul style="list-style-type: none"> • Is there a range of grammatical structures? <p><i>Grammatical accuracy</i></p> <ul style="list-style-type: none"> • Is the grammar correct? 	<ul style="list-style-type: none"> • Variety of grammatical features (tenses, structures, modals, auxiliaries, etc.) is used • Sentences and clauses are organized appropriately • Specific mistakes don't reoccur • Grammar leads to clear meaning and understanding of the ideas
<p>VOCABULARY</p>	<p><i>Lexical range</i></p> <ul style="list-style-type: none"> • Is there a range of vocabulary items? <p><i>Lexical accuracy</i></p> <ul style="list-style-type: none"> • Is the vocabulary used accurately? <p><i>Lexical relevance</i></p> <ul style="list-style-type: none"> • Is the vocabulary relevant to the topic(s) specified in the task? 	<ul style="list-style-type: none"> • Variety of words and expressions is used • Words are used accurately in terms of both meaning and spelling • The vocabulary used is relevant to the topic and text type • The words and expressions used are not completely lifted from the wording of the task

PART 3

INTERPRETING ASSESSMENT CRITERIA

3.1 Task Achievement (TA)

Task Achievement, also referred to as Task Fulfilment or Task Completion, does focus specifically on linguistic aspects of the script, but on such content and text type related features that are determined by the writing task. The complexity of the TA criterion derives therefore from the fact that whereas for example Grammar is a criterion whose definition can be considered stable over scripts written on any task type (e.g. a well-formed passive construction has the same syntactic form in a note to a friend or a covering letter), the interpretation of TA varies with each writing task. As a consequence, it is essential that raters become familiar with the concept of TA and the features of writing tasks on the basis of which they can assess TA.

Basically, with the help of TA criterion raters assess whether the script meets the text type (genres such as letters, essays, articles) and content requirements of the writing task. The features of the script that writers must pay attention to in terms of these two requirements are determined by the writing context. The three basic closely related and interdependent components of the writing context are as follows:

- **the writer**
- **the reader**
- **the writing product (the script produced by the test-taker).**

The relationship between **the writer** and **the reader** together with the purpose of the communication determine the *style* of the text: a customer (writer) writing a letter of complaint (text type) to the manager (reader) uses formal English style, whereas a person on holiday (writer) writing a postcard (text type) to a friend (reader) uses informal style. Some text types like letters or articles can be written in either formal or informal style, but other text types such as reports or research articles are strictly formal. Style may therefore be a feature inherent in the text type since it is determined primarily by the components of the writing context that are assessed under the TA criterion.

In some analytic scales designed for assessment of writing, style is evaluated together with vocabulary. However, style does not mean solely the choice of formal or informal vocabulary items – for example, the use of contracted forms and the tone of the writing also affect the style of a script. Assessing style together with vocabulary can mislead raters because it reinforces the false belief that style is determined only by the formality of words used in the script.

The script as the component of the writing context addresses aspects of a given topic and thus has a definite *content*. The script also represents a

particular *text type* (e.g. a letter, an article). Both the content and the text type are determined by the writing task and help to identify the features of the script raters must focus on when assessing TA.

Different text types have different *forms and functions*.

Letters, information brochures, compositions, articles or reports are written products that represent different *forms*. Each form has a specific *layout* (compare, e.g., the layout of a letter with that of a report) and *rhetorical organization* (the ideas included and the order they follow are different in a letter of request and in a letter of complaint).

Each text type also has a particular *purpose* which can be described either as *communicative function* (i.e. whether the text aims to describe, request, invite, give instructions, persuade) or as a *discourse mode* (narration, description, exposition, argumentation).

The constituents of the TA criterion are summarized in Figure 4 below.

Figure 4 – Constituents of the TA criterion

CONSTITUENT	FEATURES	
Style	Audience awareness Text quality	<ul style="list-style-type: none"> • formal/informal vocabulary • tone • personal/impersonal mode of expression • appropriateness • consistency
Content	Appropriate topic Relevance of ideas Coverage Length	<ul style="list-style-type: none"> • topic set in the writing task • treatment of the topic <i>development of content points</i> <i>necessary/redundant information</i> <i>range of ideas</i> • amount of detail • task solved within the set word limit
Text type	Form Communicative function Discourse mode Rhetorical organization	<ul style="list-style-type: none"> • type of written product • layout • goal of the text • typical sequencing of information

3.2 Coherence and Cohesion

In analytic scales developed for the assessment of writing the criterion called *organization* refers to the arrangement of ideas and the relationship within them in a script. Organization is an umbrella term for a number of features of a script, namely **coherence**, **unity** and **cohesion**, and as a consequence the organization criterion is also often labeled as Coherence & Cohesion.

3.2.1 Coherence

Coherence refers to the logical progression of the ideas from the beginning to the end of a script or a paragraph. In a coherent text, each paragraph

leads into another paragraph and each idea within a paragraph leads into the next idea. If a script is coherent, the reader does not have to stop and reread it in order to understand the connection between its sentences or paragraphs.

The logical arrangement of the ideas varies according to the writer's purpose, and there exist a large number of organizational patterns that can be grouped into three categories:

- those determined by the discourse mode of the script (e.g. time sequence in narration/process description, or spatial organization in a description);
- those determined by the nature of the information included in the script (e.g. order of importance: most important information first, least important last; level of specificity: information presented by moving from general-to-specific);
- those determined by the method of development (question-answer: posing a question first and providing an answer to it; problem-solution: stating the problem and proposing a solution to it).

Coherence is thus a quality of the script that derives from the relationship between the ideas expressed in its sentences and paragraphs.

3.2.2 Unity

Unity is a feature that contributes to the smooth flow of the ideas and is usually assessed together with coherence: each script develops a main topic, the paragraphs within the script develop the subtopics of the main topic, and all sentences within a unified paragraph develop the topic of the paragraph.

3.2.3 Cohesion

Cohesion is the explicit marking of the grammatical and lexical relationships between the parts of a sentence, between sentences, or between paragraphs. Cohesion can be established with several techniques of which the three that are most relevant to writing are treated below.

Reference can be realized with the use of personal or possessive pronouns, demonstratives and comparatives: it establishes a link within a sentence or between two or more sentences by connecting specific words in them.

Conjunction is another way of realizing cohesion. Co-ordinating conjunctions (*and, or, but*) connect words, phrases and clauses; adverbial connectors (*firstly, secondly, moreover, therefore, in conclusion, as a consequence*) connect clauses, sentences and paragraphs. Whereas co-ordinating conjunctions and adverbial connectors link similar linguistic units (e.g. clauses), subordinating conjunctions (*because, that, unless, when*) link independent and dependent clauses.

Lexical repetition – the use of the same word, a synonym or near synonym, or of a general word that refers to the same thing in different sentences can also establish cohesion through the creation of lexical chains that reiterate key words and concepts throughout the script.

The cohesive techniques discussed above unify the script and make it easier to read. The efficient use of these techniques makes it possible for the writer to construct sentences of varying length; to show to the reader the connection between sentences and paragraphs; and to avoid unnecessary repetition of words, sentences, or even paragraphs. The orderly arrangement and connection of ideas at sentence-, paragraph- and script-level result in a clear piece of writing. Skilful reference makes the flow of ideas easy to follow, the correct use of conjunctions indicates clearly the structure of the text, and lexical chains maintain topic continuity. The use of substitution and ellipsis results in natural and smoothly flowing text.

3.2.4 Punctuation

Punctuation in some rating scales is part of the Accuracy criterion, but it can be argued that it is more closely related to the Coherence and Cohesion criterion. In the same way that pauses and intonation help listeners to understand speech, and the layout or graphic division of a text into various functional units (addresses, titles, headings, paragraphs) indicate text structure, punctuation separates parts of a script by showing where one set of ideas ends and where the next begins, and how the parts of the script are related to one another. Without punctuation marks writing would become an incomprehensible flow of words.

A sentence without punctuation marks is difficult to interpret. Punctuation marks have clear functions, so they make it easier for readers to understand a script (e.g., the colon (:) indicates that a list or an explanation follows). However, punctuation marks not only separate a string of words but can change the meaning of a sentence. For instance, the following pairs of sentences consist of the same words and are grammatically correct, but the sentences in each pair have different meanings.

a. The woman said John is ill.

The woman, said John, is ill.

b. The Smiths have a son who is a cook.

The Smiths have a son, who is a cook.

In example *a* the first sentence means that it is John who is ill. Similarly, the comma in example *b* changes the meaning of the sentence. The sentence without the comma means that the Smiths have at least two sons, one of whom is a cook. The second sentence means that they have only one son, and he happens to be a cook.

Punctuation is directly related to meaning and if readers cannot make sense of the script because they cannot determine which words go together, where one idea finishes and the next one begins, what the relationships between sentences are, or what the exact meaning of a sentence is, they will consider the script incoherent.

3.3 Grammar

The Grammar criterion in an analytic rating scale makes it possible to assess candidate's grammatical knowledge. Raters evaluate grammatical knowledge by checking whether candidates are familiar with the form, meaning and use of a range of grammatical structures that can be expected to occur in a communication situation.

The grammar used in a script is often assessed on the basis of **range**, **degree of mastery**, and **accuracy**. It is important for raters to be familiar with what these features refer to because they may be expected to base their decisions on them when they give a score for grammar.

3.3.1 Range

Range refers to the variety of different grammatical structures used within a script: tenses, modals, passive constructions etc. A script with a large variety of grammatical structures is likely to be awarded a better score for grammar than a script that contains a few grammatical structures used repetitively. However, raters should not expect candidates to display all their grammar knowledge in a script. While some tasks require that students employ a variety of tenses, other tasks necessitate the use of modal auxiliaries or conditional sentences, so raters must adjust their expectations to the particular writing task that was used to elicit the scripts they are assessing. If the types of grammatical structures that can be expected to be found in the script are not described in detail in a rating scheme accompanying the writing task, raters should write their own script in order to see what kind of grammatical knowledge that particular task draws on.

What can be expected from students in terms of the range of grammatical structures depends not only on the writing task but also on the time and number of words set for the completion of the task. More time allows for more careful planning and redrafting, which can result in a more thoroughly constructed text and perhaps in the inclusion in the script of a larger variety of grammatical structures. The length specified for the script can also affect the variety of structures: a longer script containing developed paragraphs can be expected to display a larger variety of structures.

Consequently, raters must always interpret range by taking into consideration the topic and subtopics that must be developed in a writing task, the time allotted for the completion of the script, and the number of words set.

3.3.2 Degree of mastery

Degree of mastery refers to how familiar test-takers are with the different meanings and functions of a grammatical structure, and whether they can use it appropriately in a particular communicative situation. Candidates with varying levels of grammatical knowledge may have a different grasp

of the same grammatical structure, i.e. more proficient language learners can use one specific form to express different meanings.

Degree of mastery also refers to the students' ability to vary the form of a grammatical structure in order to express the same meaning. The construction of all grammatical structures is governed by rules. The candidates' awareness of these rules and their familiarity with the exceptions can also serve as evidence of their grammatical knowledge. Finally, a student with a good command of English can use grammar not only to express idea correctly but also to make sure that the idea is expressed appropriately in a particular situation.

Mastery thus means that the candidate can use grammar effectively and appropriately for the communication situation. Raters may decide to award points for candidates who use a range of grammatical structures even if they are not always used correctly.

3.3.3 Grammatical accuracy

Besides the range and degree of skill with which students can handle grammatical structures, accuracy is a basic feature that raters consider in the assessment of grammatical knowledge. Grammatical errors are as varied as the number of grammatical structures that exist in the English language, but a distinction is generally made between two types of errors:

- **local errors** – the ones that affect single elements within a sentence (e.g. inflections, articles, quantifiers, auxiliaries) and do not significantly interfere with the meaning of the sentence;
- **global errors** – errors that affect the whole sentence (e.g. wrong word order, missing verb or subject) to such an extent that it becomes difficult or impossible to understand it. These errors cause structural problems that render the sentences incomprehensible and thus lead to communication breakdowns in the scripts.

From the point of view of its effect on communication, it is difficult to decide without a context whether an error type constitutes a minor (local) or major (global) error. Raters must examine all the errors in their contexts so as to be able to decide to what extent they affect communication before they can label them as either minor or major errors.

The identification of faulty grammatical structures is important for the assessment of the grammatical knowledge of candidates. Raters should make sure, however, that while assessing grammatical knowledge they do not concentrate only on erroneous grammatical structures. The correct structures must also be taken into consideration: raters must weigh incorrect grammatical structures against the correct ones before awarding a mark. Unlike an interview in oral test, a script can be reread as many times as it is necessary to take a reliable inventory of correct and incorrect instances of

grammar. This does not mean that the rater must count, for example, all the correct and incorrect instances of article use in a script. Such an approach would make the rating process an extremely time-consuming activity. Instead, raters must base their judgements on the types of errors (local or global) and the ratio of correct and incorrect structures identified in the script.

3.4 Vocabulary

The Vocabulary criterion in an analytic scale is intended to assess the breadth and depth of candidates' stock of foreign words (e.g. simple words, compound ones, idioms), and of the skill with which they can handle them. Breadth refers to the number of different words they know, and depth describes their awareness of the different meanings and uses of a particular item. Knowing a word in writing means that students can use it correctly to express the exact meaning they want. For this they must know what a word refers to, what its connotation is, what other words it is related to in meaning. Moreover, candidates must be aware of how words combine in a sentence, and how they can build other words by affixation, compounding or conversion. Candidates must also be able to reproduce words correctly in writing, i.e. they must be able to spell correctly.

In general, vocabulary is assessed on the basis of **range, relevance, sophistication, and correct word choice and usage.**

In the assessment of vocabulary a distinction is made between two types of words in a script:

- **content words** – the words that fall within the category of nouns, lexical or full verbs, adjectives and adverbs;
- **function words** – such as auxiliaries, pronouns, conjunctions or prepositions – that play a mainly grammatical role.

When assessing vocabulary, raters mainly focus on content words.

3.4.1 Range

Range refers to the candidates' ability to use an adequately broad vocabulary within a script. A script that contains a large number of different words is awarded a higher score than the one in which a limited number of words and phrases are used repetitively. Candidates can avoid repetition, which is interpreted as a sign of poor vocabulary, with the use of synonyms or superordinate terms. Raters, however, must make a distinction between repetition that originates from poor vocabulary and repetition that has a well-defined function.

Range must be interpreted in relation to the task set: each task requires the development of a particular topic or subtopics, which determine the type

of vocabulary that students use. Range also depends on the required length of the script: a longer script allows for more repetition and variety

3.4.2 Relevance

Closely related to the writing task is the feature referred to as relevance of vocabulary. The writing task determines the topics and subtopics students must develop, and raters must assess whether the vocabulary used by the candidate meets the requirements of the task. These requirements may be specified in the task-specific rating scheme designed for the task and can vary from the direct identification of the nature of the expected vocabulary to such general descriptions as the type of language needed for the completion of the task (e.g. vocabulary appropriate for the expression of an opinion or formulations of a complaint). Thus, whereas raters assess whether candidates write about the topic under the Task Achievement criterion, they evaluate under the Vocabulary criterion whether students can use words relevant to the topic.

3.4.3 Lexical sophistication

Lexical sophistication refers to the ability of candidates to formulate their ideas succinctly, neatly and articulately. Students with good vocabulary skills use precise and effective words that are appropriate for the topic and with the help of which they can convey their ideas with clarity and conciseness. They can use their vocabulary to formulate complex ideas clearly, and can express themselves fluently. In order to be able to give evidence of lexical sophistication, candidates must have a broad vocabulary, must be able to express exactly what they want to communicate, and must be able to combine words so that the resulting text reads fluently and naturally. In other words, students must know a large number of words, their grammatical class, rules of word formation, and collocational patterns.

3.4.4 Correct word choice and usage

Vocabulary errors can be grouped into errors that relate to **meaning, form and spelling**.

One of the most frequent errors is the use of the wrong word to express an intended **meaning** (e.g. *journey* instead of *travel*, *learn* instead of *study*, *habits* instead of *customs*). Similarly, students often confuse two words that are closely related but opposite in meaning (e.g. *lend/borrow*, *bring/take*). Another type of meaning-related error is caused by false cognates, also called false friends. These are words that have similar forms in two languages but do not have the same meaning (e.g. *actual/current*, *sympathy/liking*).

When candidates do not know a word, they can combine existing English words to express an idea. They may use circumlocution to describe or define something if they do not know the appropriate word. Another way

students may try to make up for gaps in their vocabulary is by coining English words.

Errors that relate to **form** are caused by the candidates' inadequate mastery of word form: they are not accurate and effective enough in affixation and compounding. Besides derivation errors, students often use the wrong form of a word to express their ideas. In such cases the word does not fit into the sentence (e.g. *sport* in collocation *to do sports*).

Multi-word words (e.g. prepositional verbs, phrasal verbs, idioms) and *collocations* (words that commonly occur together) represent another source of vocabulary errors. Closely related to idioms and collocations is the type of writing error usually labeled *wrong expressions*. This is a label raters use to identify a vocabulary error that cannot be fitted into any of the above-mentioned categories: the words used in the sentence are English words that are used correctly and yet the sentence does not read naturally in English.

Spelling and **lifting** are two aspects of vocabulary that are relevant primarily to the assessment of written texts.

Spelling is especially important because incorrect spelling can impede communication. When assessing vocabulary, raters distinguish between minor and major spelling errors and check the consistency of spelling. *Minor* spelling errors do not change words to the extent that their recognition becomes difficult (e.g. *tripp* – *trip*, *sincerelly* – *sincerely*, *termal* – *thermal*). *Major* spelling errors alter a word to such an extent that it becomes very difficult to recognize (e.g. *colige* – *college*, *exercis* – *exercise*), or its spelling becomes identical with that of another word (e.g. *tripe* – *trip*, *sweat* – *sweet*, *were* – *where*, *now* – *know*). The violation of the spelling of words that receive special emphasis in teaching (e.g. *then* – *than*, *whether* – *weather*) is also considered a major spelling mistake.

Consistency in spelling is another problem that concerns the raters of written texts. Writers should use the spelling typical of one national variety of English consistently. For instance, the spellings *favour* (UK) and *favor* (US), *summarise* (UK) and *summarize* (US) should not be used in the same script.

Lifting refers to candidates copying part of the input text into their own scripts, and it becomes a problem that raters must focus on when the writing task contains input text(s). Raters cannot accurately assess the language skills of students whose scripts contain language given in the task because it is not lifting skills that they must evaluate but the candidates' ability to use their own language skills to produce continuous writing. Whereas lifting in itself affects the score awarded for the vocabulary criterion, it can also be source of vocabulary (and grammatical) errors because students are often unable to build the lifted text correctly into their own scripts.

3.4.5 Error tolerance

It is rarely the case even at the highest levels of language proficiency that raters expect candidates to produce completely error free scripts. At lower levels a larger percentage of errors and error types are tolerated. Errors that do not obscure meaning are more generally accepted than those that lead to the breakdown of communication. Nevertheless, a script with a large number of minor vocabulary errors may be awarded a lower score for the vocabulary criterion.

As in the case of the assessment of Grammar, raters must check whether it is the same vocabulary error or different errors that occur in a script. Candidates must not be repeatedly penalized for the same vocabulary error: raters must make a distinction between frequent errors and repeated instances of the same type of error.

A script that contains correct but simple vocabulary is of lower quality than a script which shows individuality in word choice and attempts more ambitious expression of ideas, even if, as a result, this is not always correct.

REFERENCES

1. Загальноєвропейські рекомендації з мовної освіти: вивчення, викладання, оцінювання. – К. : Ленвіт, 2003. – 261 с.
2. Програма з англійської мови для професійного спілкування / Г. Є. Бакаєва, О. А. Борисенко, І. І. Зуєнок та ін. – К. : Ленвіт, 2005. – 108 р.
3. Bachman L. F. Language Testing in Practice / L. F. Palmer, A. Palmer. – Oxford : Oxford University Press, 2000. – 345 p.
4. Brown H. D. Teaching by Principles: An Interactive Approach to Language Pedagogy / H. D. Brown. – White Plains, NJ : Addison, Wesley, Longman, 2001. – 314 p.
5. Fulcher G. Language Testing and Assessment: an advanced resource book / G. Fulcher & F. Davidson. – London & New York : Routledge, 2007. – 404 p.
6. Jordan R. R. Academic Writing Course / R. R. Jordan. – Harlow : Pearson Education Ltd., 2003. – 120 p.
7. Morley J. University Writing Course / J. Morley, P. Doyle, I. Pople. – Newbury : Express Publishing, 2007. – 215 p.
8. Tanko G. Into Europe: The Writing Handbook / G. Tanko. – Budapest : Teleki Laszlo Foundation, 2006. – 307 p.

Навчальне видання

ОЦІНЮВАННЯ ВМІНЬ ПИСЬМА

ASSESSING WRITING

Навчально-методичний посібник

Укладач

Гнаповська Людмила Вадимівна

Редактор *І. О. Кругляк*

Комп'ютерна верстка *Ю. М. Хижняк*

Підписано до друку 24.05.2013. Формат 60x90/16. Гарнітура Times.
Обл.-вид. арк.1,17. Умов. друк. арк. 1,5. Зам. № 1236

Видавець і виготовлювач

Державний вищий навчальний заклад

“Українська академія банківської справи Національного банку України”
вул. Петропавлівська, 57, м. Суми, 40000, Україна, тел. 0(542) 66-51-27

Свідоцтво про внесення до Державного реєстру видавців, виготівників
і розповсюджувачів видавничої продукції: серія ДК, № 3160 від 10.04.2008