

Testing Grammar: why, what and how?

Dr. Olga Kvasova

UALTA

Department for Methodology of Teaching Ukrainian and
Foreign Languages and Literatures
Institute of Philology
TSNUK



Workshop outline

Purposes of testing

Construct

Techniques

Test design



Why test grammar?

Types of tests or purposes of testing

- Standardised
(high-stakes entry or commercial tests)
- Classroom-based
(teacher-constructed)
- Summative assessment – assessment **of** learning :
achievement tests
(*mid-term, end-of year, end-of -course*)
- Formative assessment – assessment **for** learning - ‘assessment for becoming’
(John Pryor):
progress (unit) tests (quizzes)
- Placement
- Diagnostic
- Progress
- Achievement
- Proficiency

Standardised (proficiency) test: FCE (2002)

Paper 3, Use of English

- Part 1, MCQ, lexis
- Part 2, Open cloze, grammar
- Part 3, key word transformation, grammar
- Part 4, error correction, grammar
- Part 5, word-building, lexis

Total number of items – 65; grammar – 40 (61%)

Scoring (total) - 75; grammar – 50 (66%)

Time allotted – 1 hour 15'; grammar – 61-66% (?) –
49'=>

80' test paper = 64 items (?)

FCE Use of English (2002)

Part 2

For questions **16-30**, read the text below and think of the word which best fits each space. Use only one word in each space. There is an example at the beginning **(0)**.

An Egyptian crown

Two brothers have discovered that the strange looking object handed down to **(0)** *them* by their grandparents is an Egyptian treasure. It is **(16)** ... rare that is only **(17)** ... other in the world.

What skill is being tested?

Open cloze

N=15; 1 mark

- Skill tested: use of
articles,
auxiliaries,
prepositions,
pronouns,
verb tenses and forms,
phrasal verbs,
linkers

more than one possible answer -> mark scheme.

Misuse of capital letters and punctuation is ignored,
spelling must be correct.

FCE Use of English(2002)

Part 3

For question **31-40**, complete the second sentence so that it has a similar meaning to the first sentence, using the word given. **Do not change the word given.** You must use between **two** and **five** words, including the word given. Here is an example **(0)**.

Example:

0 A very friendly taxi driver drove us into town.

driven

We by a very friendly taxi driver.

What skill is being tested?

Key word transformations

N=10; up to 2 marks

- Skill tested: use of a wide range of structures, such as:

reported speech,

passive voice,

conditionals,

verb tenses,

modals +

phrasal verbs and

phrases such as *'to look forward to -ing.*

FCE Use of English(2002)

Part 4

For questions **41-55**, read the text below and look carefully at each line. Some of the lines are correct, and some have a word which should not be there.

If a line is correct, put a tick (✓) by the number **on the separate answer sheet**.

If a line has a word which should not be there, write the word **on the separate answer sheet**. There are two examples at the beginning (o and oo).

Examples: (o) ✓
(oo) of

What skill is being tested?



Error correction

N=15; 1 mark

Skill tested:

recognition of incorrect verb forms,
wrong pronouns,
prepositions and
articles.

The absence or misuse of capital letters is ignored, although spelling must be correct.

FCE (2002)

OPEN CLOZE

Use of articles,
auxiliaries,
prepositions,
pronouns,
verb tenses
and forms,
phrasal verbs,
linkers

KEY WORD TRANSFORMATION (PARAPHRASE)

Use of articles, auxiliaries,
prepositions, pronouns,
verb tenses and forms,
phrasal verbs, linkers

ERROR CORRECTION

Use of articles,
auxiliaries,
prepositions,
pronouns, verb
tenses and
forms, phrasal
verbs, linkers

TEST SPECIFICATIOIS related to CEFR

Is a proficiency test truly representative?

“.. we cannot be sure confident that a (proficiency) grammar test includes a good sample of all possible grammatical elements. But the very fact that there can be **so many** items does put the grammar test at an advantage.”

Hughes (1996:141)

“.. there is often good cause to include a grammar component in the achievement, placement and diagnostic tests of teaching institutions.”

How do we prepare a grammar test?

- Use a ready-made standardised test (FCE, etc.) or progress/achievement test from the course book set (Global, etc.)
- Use any ready-made test as a base and modify it (pick out only relevant items, modify difficulty, combine with own-constructed items)
- Develop own test

Any other way? Which exactly?

Which items types do you love using?

Survey of grammar test preparation practices

Most frequently used item types

- 1 Alternative choice + gap-filling
- 2 MCQ + MCQ (theory)
- 3 Paraphrase
- 4 Matching functions
- 5 Error correction

Most difficult items:

to develop (as teachers)

- 1 MCQ
- 2 Matching functions
- 3 Error correction
- 4-5 **Paraphrase** +MCQ
(theory)
- 6 **Gap-filling**
- 7 **Alternative choice**

to do (as students)

- 1 **Paraphrase**
- 2 Error correction
- 3 Matching functions
- 4 **Gap-filling**
- 5 MCQ (theory)
- 6 **Alternative choice**
- 7 MCQ



Do we (ever) follow the testing cycle?

- Design test specifications
- Develop a test within a team
- Pre-test (trial and pilot) items with the provided keys
- Analyse feedback
- Modify the test

WHY? WHY NOT?

How to develop a grammar test

1 Writing specifications

What and how to test

(structures, text-based/discrete, item types, number of items, time allotment, scoring)

“For achievement tests where teaching objectives or the syllabus list the grammatical structures to be taught, specifications of content should be **straightforward**. Where there is no such listing it becomes necessary to **infer** from textbooks and other teaching materials what structures are being taught.”

Hughes (1996)



2 Sampling

- Selecting widely from structures specified
- Regarding what structures are most or least important in the teaching context
- Full content representation vs. content under-representation
- Test should not deliberately concentrate on the structures which happen to be easiest to test (Hughes)

Outcome: content validity of a test

Syllabus for Grammar (excerpt)

- Infinitive as a means of denoting secondary actions/ secondary predication.
- Grammatical meaning, categories and cognitive implications of the infinitive.
- Forms and kinds of the infinitive: grammatical meaning and usage.
- Syntactical vs rhetorical functions of the infinitive.
- Secondary predication constructions with the infinitive: rhetorical purpose and usage. ->

Should it be further specified? In which way?

3 Writing items

We need to consider:

- Sentence level or text level?
- Recognition, recognition & production, production?
- What kind of response do you plan to receive? Why?
- Which item types /techniques to use?
- How to keep the right ratio between varied difficulty and importance of items in term of their weighting?
- How many tasks/items to include?
- How much time to allot?

Do you take the above into consideration while designing your test paper? Discuss these issues with you partner(s).

Task types based on expected response formats (Purpura)

	Constructed-response data
Selected-response data	
<ul style="list-style-type: none">•MC•true/false•same/different•matching•gram. judgments	
Useful for testing receptive knowledge ; assessing for intake	

Constructed-response data

Selected-response data

- MC
- true/false
- same/different
- matching
- gram. judgments

Useful for testing **receptive knowledge**; assessing for **intake**

Limited-production data

- gap-fill
- cloze
- sentence completion
- DCT
- short answer

Useful for testing **emergent knowledge**; assessment for **system development**

		Constructed –response data		
Selected-response d	Limited-production	Extended-production		
<ul style="list-style-type: none"> •MC •true/false •same/different •matching •gram. judgments 	<ul style="list-style-type: none"> •gap-fill •cloze •sentence completion •DCT •short answer 	product-focused	performance-focused	process-focused
		<ul style="list-style-type: none"> •essay •interview •reports •projects •presentation 	<ul style="list-style-type: none"> •<i>Simulations</i> (role-plays) •<i>Recasts</i> (retellings) •<i>Practice</i> (decision-making) 	<ul style="list-style-type: none"> •<i>Observation</i> (checklists) •<i>Discussion</i> (problem solving) •<i>Reflections</i> (journals, logs)
Useful for testing receptive knowledge ; assessing for intake	Useful for testing emergent knowledge ; assessment for system development	Useful for testing full production ; Assessing for output		

4 Scoring constructed-response tasks

“The important thing is to be clear about what each item is testing, and to award points for that only. There may be just **one element**, such as the definite article, and all available points should be awarded for that; **nothing should be deducted for non-grammatical errors**, or for errors in grammar which is not being tested by the item. For instance, a candidate should not be penalised for a missing third person *-s* when the item is testing relative pronouns; *opend* should be accepted for opened, without penalty.”

Hughes (1996:145)

Scoring selected- response task

The **key** – unambiguous, including all possible correct answers

Constructing test items: selected-response

- MCQ
- Hughes and other authors advise **to avoid excessive use of MCQ:**
- It is difficult to find suitable **distractors** due to the **limited paradigms** existing in English even in case of 3 options of response
- It is **undesirable to 'invent'** non-existing forms to use as distractors =>
- it is more reasonable and convenient to use **alternative choice. Responses are circled or indicated A or B. Key is required**

Constructing test items: limited-production-response

- Modified cloze

Originally, every 9th word is deleted.

MC : testing functional language (use of articles, adverbs, prepositions and linking words).

Responses are written/produced where indicated

- Gap-filling

Open the brackets; supply the correct form of the verb/adjective/adverb.

Key is required, sometimes with two correct responses.



Paraphrase (key word transformation)

- Each question contains three parts:
 - 1) a lead-in sentence,
 - 2) a key word, and
 - 3) a second sentence of which only the beginning and/or end are given.
- It is used to test the knowledge and feel for structures.
- **A careful preparation of key is required**

Extended-production response

- **Translation** – in monolingual classes.
- We test **grammatical material** including spelling → vocabulary should be high frequency to provide the purity of items.
- **Essay, report** etc.:
- Holistic scoring (one score for several dimensions 0-5)
- Analytic scoring (one score for each dimensions 0-5 x 3 =)

Developing achievement test specifications: what and how?

Quiz (short progress test)

- What to test:
- Part of the content (covered)
- Discrete elements of the content are tested
- Focus on receptive knowledge, discrimination and use of forms
- Number of items and time allotment are limited
- How to test:
- Selected response (mostly)
- Limited-production and extended-production response limited to the skills developed so far

Mid-term; end-of-term paper

- What to test:
- The entire content should be represented evenly
- Focus on emergent knowledge and full production → assessing for output
- Number of items (50) and time allotted (75')
- How to test:
- Limited use of selected response
- Wide use of limited- and production response items



Writing items

I Selected- response

Multiple/alternative choice (1 mark)

II Limited-production-response

Gap-filling (1 mark)

Paraphrase (2 marks)

III Production response

Translation (? marks)

Essay (analytical scale)

N of items as well as timing vary




Developing a quiz

- **What to test:**
- Part of the content (covered)
- Discrete elements of the content are tested
- Focus on receptive knowledge, discrimination and use of forms
- Number of items and time allotment are limited
- **How to test:**
- Selected response (mostly)
- Limited-production and extended-production response limited to the skills developed so far



Task 1

- **1.1. Write a selected response item**
- **1.2. Write a limited response item**
 - **Use the same original sentences**

- 
- 1 He is pleased to have met you at the party. He's been talking about your meeting for two days.
 - 2 You seem to be working harder than usual this month. You sound much more confident.
 - 3 I would like to have seen his face when he learnt about his newborn baby.
 - 4 I'm sorry to have mentioned your name during the conversation. I noticed you frown at that moment.
 - 5 It was unreasonable of them to have been chatting during the entire lecture. It got on the lecturer's nerves.

Key 1.1. Alternative choice (selected response).

1 Circle the correct form of the infinitive.

1 He is pleased **to meet/to have met** you at the party.
He's been talking about your meeting for two days.

2 You seem **to work/to be working** harder than usual
this month. You sound much more confident.

3 I would like **to see/to have seen** his face when he
learnt about his newborn baby.

4 I'm sorry **to be mentioning/have mentioned** your
name during the conversation. I noticed you frown at
that moment.

5 It was unreasonable of them **to have chatted/to have
been chatting** during the entire lecture. It got on the
lecturer's nerves.

Key 2.1. Gap filling (limited-production response)

Open the bracket supplying the correct form of the infinitive.

- 1 He is pleased to ... **(meet)** you at the party. He's been talking about your meeting for two days.
- 2 You seem to **(work)** harder than usual this month. You sound much more confident.
- 3 I would like to **(see)** his face when he learnt about his newborn baby.
- 4 I'm sorry to ... **(mention)** your name during the conversation. I noticed you frown at that moment.
- 5 It was unreasonable of them to ... **(chat)** during the entire lecture. It got on the lecturer's nerves.

- Taks 3: Write a limited production response item. Discuss what skill is tested and which format to choose.
- 1 She felt that her hands trembled.
- 2 I was rude to him, which was stupid.
- 3 It is necessary that everyone should know the truth.
- 4 I don't expect you'll think of me.
- 5 Did you have any difficulty in dealing with the correspondence?
- 6 As a student I usually swept the floor in my room.
- 7 It seems that he was very rich in the past.
- 8 They say this girl is a celebrity.
- 9 He had to wait a long time until his paper was published.
- 10 She is clever; she won't quarrel with her mother-in-law.

Key3. Paraphrase (limited-production response).

1 She felt that her hands trembled.

... her hands

2 I was rude to him, which was stupid.

It was**to him.**

3 It is necessary that everyone should know the truth.

... for.....**to know the truth .**

4 I don't expect you'll think of me.

... to think**of me.**

5 Did you have any difficulty in dealing with the correspondence?

Was it difficult**correspondence?**



Translation task - do it yourself!



Mid-term paper

Your suggestions:



Thank you for participation!

-
- olga.kvasova.1610@gmail.com
ualta@ukr.net
 - Visit our website: ualta.in.ua

UALTA

Ukrainian Association for Language
Testing and Assessment