# Training university teachers in developing and implementing rating scales to assess L2 writing

**Olga Kvasova**

**Taras Shevchenko University of Kyiv**

**UALTA**

# How well can teachers assess L2 writing?

- Increased needs in EL writing

- Absence of clear standards to assure quality of writing in HEIs

- Lack of teacher training in assessing writing →

- What and how is writing assessed today?

- What are the gaps in AL in writing and how to bridge them?

# Survey of LAL in assessment of writing in universities

- January-February 2018
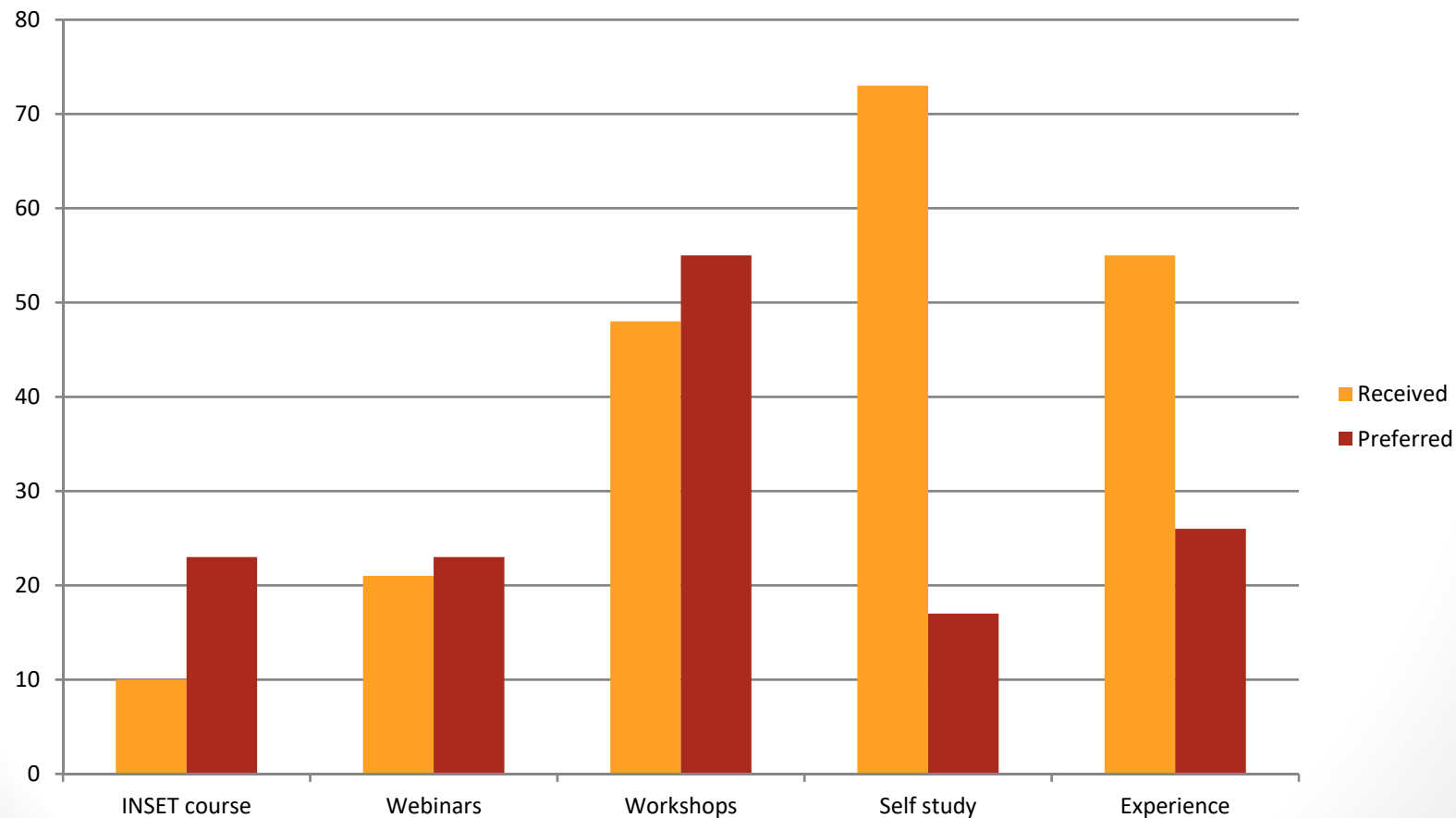- Respondents:104

ESP teachers – 68

Teachers of General English (linguistic/teacher-training universities) – 36
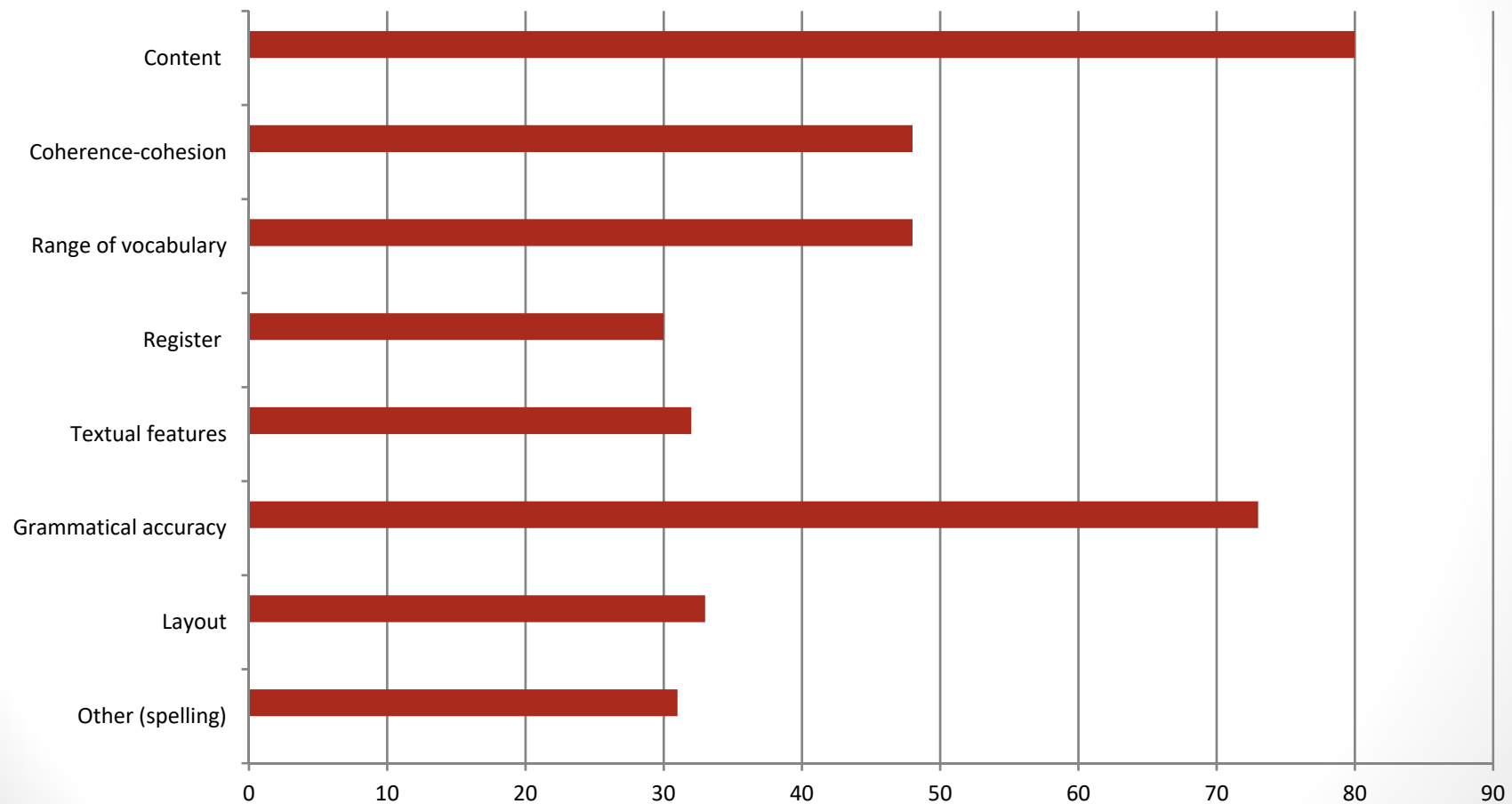
- Survey: 12 questions

Current practices

Training received and needed

# Training in assessment of writing: received & preferred

# Criteria for assessment of writing

# Current study

- Training of three raters at CRELLA; designing research
- Pre-training preparation: expert raters' (ER) use of initial scale (10 adverts), discussing→ explaining in writing why they awarded each score.
- Training: presentation and explanation of the initial scale; mock rating (10 adverts); comparing own scores with ERs' scores; discussing →any change of opinion after discussion?
- Rating (independent) of 100 students' papers (letters of complaint)
- Filling in Questionnaire 1 (on use of each of the criteria while rating)
- Filling in Questionnaire 2 (on use of criteria after rating)
- Using spreadsheets to compare and contrast scores

# Initial rating scale

| Marks | Textual features (TF) | Coherence (COH) & cohesion (coh) | Vocabulary (V) & register (R) | Grammar |
|---|---|---|---|---|
| **9-10** | Meets all text types requirements (TTR) | *Fully COH text; coh on sentence and paragraph level* | Wide range of V, correct choice of words in compliance with R | *Wide range of structures relevant to TF, few minor inaccuracies* |
| **7-8** | Meets major TTR | *COH text; appropriate sentence and paragraph-level coh* | Good range of V with few cases of wrong choice of words (WCW); few inconsistencies in R | *Good range of structures relevant to TF, some inaccuracies not hindering COMM* |
| **5-6** | Frequent inconsistencies in meeting TTR | *Sentence-level coh noticeable, lack of paragraph-level coh* | Limited range of V with frequent cases of WCW; frequent inconsistencies in R | *Limited range of structures, frequent inaccuracies hindering COMM* |
| **0 - 4** | | | | |

# A letter of complaint



You have bought a product advertised by an online store like this:

**Ung Gear S3**
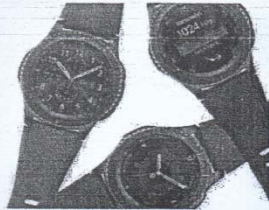
⭐ 5 STARS

Ung's latest smartwatch has a 1.3″ AMOLED touchscreen, a built-in speaker and a microphone for calls, and a battery that lasts up to 4 days per charge. It's also compatible with standard watch straps. The Frontier model has 3G/LTE connectivity.

In reality, you were frustrated with its quality, delayed delivery and poor service.

Write a letter of complaint to the online store in no more than 120 words.

Dear Sir / Madam

I am writing to express my dissatisfaction with the quality of smartwatch, which I bought in your online store 3 days ago. On using this smartwatch for the first time, I found out that the microphone for incoming calls is faulty, it is not possible to answer for calls.

Also, there is a problem with a battery of smartwatch. It is said that the battery lasts up to 4 days per charge but it lasted only 1 days without charge.

In addition, the delivery of the product was delayed. I was promised to recieve my product on Monday 5th of October, but in reality I have received my parcel only on Thursday 8th of October.

To resolve the problem, I would appreciate if you refund my money. Enclosed are copies of my receipts and guarantee.

I am looking forward to hearing from you about the resolution to my problem.

Sincerely,

Jane Ostin,

# 1. Textual features

**Check:** is the script written in compliance with the **text type requirements**?
**Look for:**

| communicative purpose | □ fully achieved | □ mostly achieved | □ partially achieved | □ not achieved |
|---|---|---|---|---|
| *standard composition* | *□ fully observed* | *□ mostly observed* | *□ partially observed* | *□ ignored* |
| **rhetorical functions of each part of script** | □ fully observed | □ mostly observed | □ partially observed | □ ignored |
| *clarity* | *□ totally clear* | | *□ not quite clear* | *□ unclear* |
| **register** | □ observed | | □ partially observed | □ ignored |
| *tone* | *□ relevant* | | *□ not quite relevant* | *□ irrelevant* |
| **layout** | □ relevant | | □ partially irrelevant | □ irrelevant |
| *length* | *□ relevant* | | *□ not quite relevant* | *□ irrelevant* |

# Textual features typical of <u>letter of complaint:</u>

- Is the **composition standard** for this type of text?
- Is there **an address** and a proper **salutation**?
- Is the **purpose of writing clearly stated** in the *opening paragraph*?
- Does the *body contain* a **summary of the events** that prompted the complaint? Does it include **relevant information** (exact dates, time, location of purchase, price, inconvenience caused)?
- Does the *body* contain the **steps taken to resolve** the problem (communication with the company)?
- Does the *final body* paragraph state **expectations** of how the company should resolve the matter, including specific actions and deadlines?
- Is the letter **focused, clear and logical**?
- Is it **formal** as addressing an unfamiliar person?
- Is the **tone calm, gracious** (without personal attacks and abusive language)?
- Is the **layout relevant** to the text type conventions?
- Is the letter of appropriate length (quite **short)**?

# Questionnare 1 (while rating)

| Ex. | It was **overall**<br>☐ easy<br>☒ not quite easy<br>☐ quite difficult<br>☐ difficult to rate | It was **quite convenient** to rate<br>☐ Textual features (TF)<br>☒ Coherence-cohesion (CC)<br>☐ Vocabulary & register (VR)<br>☒ Grammar (Gr) | It was **rathter problematic** to rate<br>☒ Textual features (TF)<br>☐ Coherence-cohesion (CC)<br>☒ Vocabulary & register (VR)<br>☐ Grammar (Gr) |
|---|---|---|---|
| VI. 01 | ☐ easy<br>☐ not quite easy<br>☐ quite difficult<br>☐ very difficult | ☐ TF ☐ CC ☐ VR ☐ Gr | ☐ TF ☐ CC ☐ VR ☐ Gr |
| VI. 02 | ☐ easy<br>☐ not quite easy<br>☐ quite difficult<br>☐ very difficult | ☐ TF ☐ CC ☐ VR ☐ Gr | ☐ TF ☐ CC ☐ VR ☐ Gr |

# Questionnaire 2 (after rating)

| # | Statement | Fully agree | Quite agree | Quite disagree | Fully disagree |
|---|-----------|-------------|-------------|----------------|----------------|
| 1 | The scale is rater-friendly and easy to use | | | | |
| 2 | The scale is comprehensive (it considers all relevant aspects of writing) | | | | |
| 3 | The criteria are appropriately grained | | | | |
| 4 | The weighting is fair | | | | |
| 5 | The scale seems a reliable tool to assess writing | | | | |

# (Q2 cntd) Comments on the scale

1.The scale misses some important aspects of writing, such as … . I suggest introducing such criteria as …. .

2. Criteria *…* should be grouped together.

3. Criteria *…* should be split.

4. Criteria *…* should be weighted differently:  for instance,… .

5. Criteria… are formulated in a confusing way. I suggest reformulating them as … .

6. The biggest problem that I faced while rating was … . I suggest  ….

7. My overall evaluation of the scale is … .

# Interim results: Quare 1

- **Rating procedure:**

Easy                          – **65%**

Not quite easy – **26%**

Quite difficult  –  **8%**

Very difficult    -    **0**

- **Rater-friendly criteria** vs **Confusing criteria**

                        **77%** vs **23%**

   **TF  and C&C**                    **V&R and GR**

# Interim results: Quare 2

**R3:** "The biggest problem that I faced while rating was *to decide whether the writing meets major text type conventions of the Textual Features criterion. It was really confusing to go through all the conventions such as pragmatic purpose, standard composition, clarity, register, length, tone every time while rating Textual features. Often I could not decide how many points I had to subtract if a writing didn't comply with one / two or more text type conventions. As a result I found the criterion Textual Features too general.*"

**R7:** *"My overall evaluation of the scale is positive as it is comprehensive and includes all the aspects of writing. Some of the criteria were confusing for me as they included a wide range of conventions. Also, in my case I felt I lacked training, and sometimes I needed support to double check if my reasoning was right.*"

# Interim results: Quare 2

**R 8:** "Criteria Textual Features, Cohesion and Coherence should be weighed differently:  for instance, *each of them should be allotted 2 points maximum. It will simplify the rating process. For example, if a task fully meets the requirements – then a rater gives 2 points, partially meets the requirements – 1 point, and does not meet – 0.*"

**R 10: "***The scale does not miss any aspects of writing. But some Criteria, for example Textual Features, include too many conventions, which complicated the process of rating. I suggest introducing Organisation as a separate criterion.*"

**R3:** "*Another criterion I want to discuss is Grammar. I did not know how to deal with frequent grammar inaccuracies which did not hinder my understanding of the pragmatic purpose: did I have to give a lower score or not?*".

# New criteria

- Task achievement

- Genre conventions

- Coherence & cohesion

- Vocabulary

- Grammar

olga.kvasova.1610@gmail.com

ualta@ukr.net

# THANK YOU!