TEACHING AVIATION WEATHER VOCABULARY TO ROMANIAN AIR FORCE CADETS – QUANTITATIVE ERROR ANALYSIS BASED ON TRADITIONAL MATCHING AND MULTIPLE-CHOICE EXERCISES

Ramona HĂRȘAN, Cosmina-Oana ROMAN, Kinga KOLUMBÁN

"Henri Coandă" Air Force Academy, Braşov, Romania (ramona.harsan@afahc.ro, roman.cosmina@afahc.ro, kinga.kolumban@afahc.ro)

DOI: 10.19062/1842-9238.2024.22.2.10

Abstract: *The main focus of the article is on the quantitative proportion and distribution of aeronautical meteorology adult learner errors around a few main vocabulary topics in aviation weather as ESP: precipitation (liquid, freezing, frozen), the atmosphere (air, sky cover, visibility, air movement, cloud formations) and their effects (runway conditions). The purpose is to analyze errors found in undergraduate (Air Force cadet) classroom-produced materials featuring rather traditional vocabulary-focused activity types (matching and multiple choice lexical reinforcement exercises) in order to identify areas of difficulty and lay the groundwork for further quantitative and qualitative approaches on the specific challenges affecting Romanian learners studying aviation English as EPP.*

Keywords: *metacognition, ESP/EPP, aviation weather vocabulary, error analysis, error distribution*

1. INTRODUCTION: THEORETICAL BACKGROUND AND PURPOSE

It is generally thought that learning about the general grammar and vocabulary of a foreign language is sufficient for second language (L2) acquisition (SLA) – and of course, it remains a prerequisite. However, one of the many new lines of thought in SLA-related research revolves around the general concept of metacognition, which refers to a learner's induced self-awareness in understanding the processes articulating their own acquisition of knowledge. This approach to teaching and learning has opened a new and engaging territory for educational studies in both science and humanities, from mathematics to Foreign Language Teaching (FLT) and methods in English as a Second Language (ESL) teaching [1, 2, 3]. Among the main lines of attack, error analysis is seen as one of the cornerstones of student/adult learner ability enhancement [4, 5], providing such learners with the necessary vital instruments to assist them in critically assessing their own performance and successful or unsuccessful learning habits and mental processes. There are also researchers who have implied that metacognition and error analysis was to be integrated into their respective L2 learning frameworks [6].

To add to this line of thought, the present article is a first step in a series of researches meant to analyze errors: 1. from the perspective of their quantitative distribution around common topics in Aviation English as a college/university/higher education study subject, and 2. the subsequent qualitative interpretation of the typical mistakes identified under stage 1, focused on a perspective meant to highlight the bearing of L1 interference

in FLA, to the ultimate aim of covering the basic difficulties that Romanian users of ESL as EPP in the aviation industry come across in their learning process. In the long run, this identification and interpretation of mistakes is to support the potential metacognitive enhancement of linguistic abilities with this specific group of learners.

Note should also be made that we consider this contribution to be a pre-research experiment done within the framework and set timeline of a research project meant to support further analyses (see Acknowledgement), to be continued and developed as new generations of cadets agree to furnish additional material.

As such, this article focuses in particular on aviation weather vocabulary and its main difficulties as they result from the quantity and topic-centered organization of errors in Romanian Air Force cadet classroom-produced worksheets and final exam simulation tests, with the objective of assisting the educational process in typical error identification.

The methodology is based on seminal contemporary literature in the field [7, 8].

2. METHODOLOGY

For this limited experiment, four mixed-ability undergraduate student groups' classroom produced materials were subjected to analysis. The learners' approval for their work to be used under the protection of anonymity was asked and received at the beginning of the respective courses. The four military cadet groups involved belonged to two successive classes counting 20, 12, and 10 undergraduates, respectively. The largest of the three was made up by two compeer air traffic management-oriented groups (from now on called A and B for privacy concerns), while the other two were two successive meteorology-oriented groups (the largest of the two will be codenamed C, while the smallest will be assigned the letter D).

For the scope of this article, specialist vocabulary proficiency was tested by means of a small variety of traditional exercise types, involving limited, standardized language contexts such as the (as-such or slightly rephrased) definitions in matching exercises (upper intermediate-level structures, used as first semester material), with the objective of noticing both the quantitative bearing of the mistakes and their proportional distribution around a few main semantic fields organized as course topics: precipitation (liquid, freezing, frozen), the atmosphere (air, sky cover, visibility, air movement, cloud formations) and their effects (runway conditions). Specifically, the activities were: two intermediate-level matching exercises (final test simulation reinforcement exercises) and a vocabulary reinforcement multiple-choice exercise.

Student work was done on paperback handout materials, collected and processed within the Academy's Foreign languages dedicated lab (supported and financed through the research project mentioned in the „Acknowledgement" section at the end of the article), along with corresponding Professor feedback sheets. The full scanned or paperback material is not included here for obvious reasons (space concerns), but all the errors in the feedback materials containing personalized feedback for each individual student's work were collected manually under tabular form with no exception whatsoever, to assure the correctness and reliability of the analytical approach.

To calculate percentages and do the statistics, rather simple in this pilot study due to the dimensions and low degree of complexity of the corpus involved, no specific software was used, except for a few researcher-supervised data syntheses (calculations) provided by Open AI's free version of Chat GPT. All mistakes were manually introduced in the specific tables mentioned above (and analyzed in the following sections) for each of the tasks. In order to analyze them from the perspective(s) we were interested in, heuristic methods were used to interpret the above-mentioned tables and synthesize conclusions.

3. CORPUS ANALYSIS

As mentioned under section „2. Methodology", the actual corpus consists of 39 student written contributions and the corresponding 39 one-on-one Professor feedback sheets, distributed as follows:

- 29 responses/answer sheets (8 A +10 B + 11 C) to two vocabulary-centered matching exercises, out of which the first activity was separated into 3 topic-oriented tasks (I. Precipitation, II. Snow, III. Air, wind, visibility and clouds) and managed simultaneously in 3 laboratory sub-groups (via specific row seat allocation to avoid copying); the other was a mixt exercise synthesizing essential vocabulary loosely organized around the same topics, distributed simultaneously within group C; out of the 20 undergraduates in the first group (AB), 2 were absent at the time the assignment was given, while 11 of the 12 group C cadets were present;
- 10 responses/answer sheets (10 D) to an intermediate-level classic multiple-choice vocabulary exercise (10 out of 10 cadets were present in class as the task was given);

The distribution above resulted in 4 „error-collection" tables: 4 error-collection tables for the two vocabulary-centered matching exercises considered together (3+1, respectively, according to the specific row/group allotment) and 1 error-collection table for the intermediate-level multiple-choice vocabulary exercise.

Thus, 9 to 11 answer sheets coming from four different workgroups, two specializing in air traffic management (ATM), the other two specializing in aeronautical meteorology (AM) were collected for each specific exercise (9 to matching I + matching III, 9 to matching II, 11 to the matching vocabulary mix and 10 to the multiple-choice exercises), while all cadets were of perfectly similar age groups, qualifications and years of study. We also chose to specifically compare study profiles with slightly different focuses on aviation weather English as a subject for learning: while it is one of the important thematic concerns of the ATM-oriented English course, if forms the ultra-specialized core of the contents designed for the AM-oriented English course. The total amount of answers received from the cadets in all the four groups (ABCD) was 271.

The following analysis will detail the errors in the sense of error quantification, topic-related distribution and proportions, according to the criteria set in the previous sections (see *1. Introduction: Theoretical Background and Purpose*).

3.1 Two matching exercises. As a brief description of the exercises, the „I. Precipitation" task was to match 5 words and phrases and their definitions. The specific terms to be defined, displayed in a vocabulary box framing a string of concepts separated by bullets were shown as: 1. FZRA (freezing rain), 2. PL (ice pellets), 3. GS (graupel/soft hail), 4. GR (hail), 5. build-up. Correspondingly, 5 short, mildly rephrased definitions (as compared to the definitions present in course materials) were provided. The correct order of the concepts after matching the definitions was 2,1,3,4,5. The same organization (5 words and phrases and 5 mildly rephrased definitions were provided and displayed in the same manner) for the „II. Snow" and „III. Air, wind, visibility and clouds"-related tasks. Under task „II. Snow", the concepts to be defined were: „1. snow rut, 2. snow drift, 3. compacted snow, 4. snow bank, 5. build-up", and the key was 1,2,5,3,4, whereas under task „III. Air, wind, visibility and clouds", the concepts were displayed as „ 1. crosswind, 2. wind shear, 3. gust, 4. BKN, 5. SCT" (with no expansion of the sky-cover descriptor acronyms, as they were considered to have been all-present during class activities throughout the semester); the key read 2,1,3,4,5. These shorter, thematic exercises were used to train the ATM groups (designated here as A and C), for whom the final exam was meant to include this type of simpler, topic-specific matching exercise to check the aviation weather jargon.

The „Vocabulary Mix" tasks in the final simulation tests applied during a final exam simulation test including 11 concepts: 1.GR, 2.wind shear, 3.(cloud) tops, 4.cloud ceiling, 5.moderate chop, 6.backing wind, 7.veering wind, 8.snow rut, 9.GS, 10.crosswind, 11.snow drift (with 6, 5, 2,7,10, 3, 4, 8, 11, 9, 1 as key). Of course, this more complicated test design was used to train the aviation weather group (designated as C) for their final exam matching exercise.

Student answers turned out as described in Tables 1-5, where each cadet was assigned a number and a group designator (A and B for the two lab groups studying air traffic management, C for the group studying aviation weather as the core of their respective curricula), to observe the privacy-related ethical concerns mentioned under section „2. Methodology". Asterisks in the first four tables mark the fact that one wrong answer implied the other (the same confusion between terms resulted in two wrong answers at once).

Table 1. I. Precipitation

	Std. code	*No. of errors/ No. of answers*	*Details (based on answer key: 2,1,3,4,5)*
1.	A2	0/5	–
2.	A5	0/5	–
3.	A8	4/5	Answer: 4,2,1,3,*5* • *GR* identified as *PL* • *PL* identified as *FZRA* • *FZRA* identified as *GS* • *GS* identified as *GR*
4.	B10	2/5	Answer: 4,*1*,*3*,2,*5* • *GR* identified as *PL** • *PL* identified as *GR**
5.	B9	4/5	Answer: 4,2,1,3,*5* • *GR* identified as *PL* • *PL* identified as *FZRA* • *FZRA* identified as *GS* • *GS* identified as *GR*

This is an assessment of the 25 answers listed under the 5 worksheets received.

According to the 5 answer sheets under consideration, 2 cadets out of 5 were able to perform flawlessly, while the other 3 made mistakes. Among the cadets in the first group having performed the same activity (designated as group A), 2 out of 3 respondents answered perfectly (A2 and A5), while the third (A8) was only able to come up with one correct match. In the second group (designated as group B), both respondents made mistakes: B10 made 2 mistakes, while B9 made 4 out of 5. It may also be relevant that nobody mistook concept 5 (build-up) for any of the other terms or phrases, and that most mistakes involved, from the highest to the lowest number: the confusion between *4.GR* and *2.PL* (3+1), then with an equal number of occurrences, *3.GS* and *4.GR* (2), *2.PL* and *1.FZRA* (2), *1.FZRA* and *3.GS* (2). It may also be relevant to notice that A8 and B9, even though belonging to two different generations, made the exact same mistakes. We also need to note that B10's mistakes are based on a single confusion (between *4.GR* and *2.PL*).

The proficiency level of this particular sub-group seems to be by far the lowest in AB, judging by the fact that a majority of 60% (3 out of 5) of the cadets making at least two mistakes (a reciprocal mismatch or two different wrong answers) and 40% of the entire number of answers given counted as wrong.

We must also note that out of the 10 wrong answers given, only one was a reciprocal mismatch, accounting for only 2 incorrect solutions (20%). Additionally, the number of mistakes per student is also the highest (2 learners missed out on 4 matches, and this with no reciprocals in any of the cases, against only a maximum of two per student in the other 2 workgroups, all generated by mutually implicating errors resulting in 2 wrong answers at once).

Table 2. II. Snow

	Std. code	*No. of errors/ No. of answers*	*Details (based on answer key: 1,2,5,3,4)*
1.	A1	0/5	–
2.	A3	2/5	Answer: *1*,*2*,3,5,*4* • *snow bank* identified as *compacted snow** • *compacted snow* identified as *snow bank**
3.	B2	0/5	–
4.	B3	0/5	–
5.	B4	0/5	–
6.	B5	0/5	–
7.	B6	0/5	–
8.	B7	0/5	–
9.	B8	0/5	–

This is an assessment of the 45 answers received on the 9 worksheets.

A very visible result here is the fact that all cadets in group B answered the grid perfectly, while among the two cadets in group A who received this assignment, one was just as proficient. Therefore, out of the total of 9 respondents, only one made mistakes (A3), and those mistakes involved a reciprocal confusion: as the phrase *5.snow bank* was mistaken for *3.compacted snow*, the first mistake implied the second. Extremely high proficiency seems to characterize about 88.88% of the group, with approximately 95.55% correct matches out of the total number of answers given.

Table 3. III. Air, wind and visibility and clouds

	Std. code	*No. of errors/ No. of answers*	*Details (based on answer key: 2,1,3,4,5)*
1.	A4	0/5	–
2.	A6	0/5	–
3.	A7	0/5	–
4.	B1	2/5	Answer: *2*,*1*,*3*,5,4 • *BKN* identified as *SCT** • *SCT* identified as *BKN**

This is an assessment of the 20 answers received on 4 worksheets.

In this case, 3 out of 4 respondents made all matches correctly, and all 3 were part of group A. B1's mistake also only involved a reciprocal confusion: as the phrase *4.BKN* was mistaken for *5.SCT* and the other way around. This means 75% of these cadets' performances are indicative of extremely high proficiency, with only 10% errors among all the answers given, while the cadet with the lowest ability still answered 60% of their test correctly.

Table 4. Vocabulary Mix

	Std. code	*No. of errors/ No. of answers*	*Details (based on answer key: 6,5,2,7,10,3,4,8,11, 9.1)*
1.	C1	0/5	–
2.	C2	0/5	–
3.	C3	0/5	–
4.	C4	0/5	–
5.	C5	0/5	–
6.	C6	0/5	–
7.	C7	0/5	–
8.	C8		Answer: *6,5,2,7,10,3,4,8,11*,1,9 • *GR* identified as *GS** • *GS identified as GR**
9.	C9		Answer: *6,5,2,7,10,3,4,8,11*,1,9 • *GR* identified as *GS** • *GS identified as GR**
10.	C10		Answer: *6,5,2,7,10,3,4*,11,8,*1,9* • *snow rut* identified as *snow drift** • *snow drift* identified as *snow rut**
11.	C11	0/5	–

In group C, 121 answers received via the 11 worksheets considered for assessment. 8 out of the total of 11 respondents made all matches correctly, with only 3 committing errors, which brings us to an approximate 72.72% vs. 27.27% score in favor of extremely high versus high learner proficiency and a total of 4.95% vs 95.04% proportion of erroneous vs. right answers. It must also be noted that all six errors are reciprocal (i.e., based on single confusions resulting in two simultaneous mistakes), and that both C9 and C10 made the same mistake – mistaking *1.GR* for *9.GS* – while the other mistake involved a confusion between *11.snow rut* and *8.snow drift*. Note should be made that just like with groups A and B, the vocabulary revolving around air, air movement, wind, visibility and clouds was the least problematic (as it posed virtually no challenge at all).

By comparison, groups AB scored generally higher, but with a low in the semantic field of precipitation, leading to a mean result of 20 out of 90 answers gone wrong (an error percentage of 22.22% vs. 88.88%, approximately) and 8 cadets out of 18 committing at least a reciprocal (44.44%).

The rather visible difference could be accounted for by the fact that group C studied aviation weather as one of its core subjects, while AB groups didn't.

There is also the larger number of learners and answers, but while the first is relatively relevant, let us also remember that the second is not: while in group C each cadet had to do an 11-match exercise, in groups AB each learner only needed to execute 5 matches correctly. The time management should be counted as virtually identical, as during the same test, cadets in group C had 5 other matches to do (in ATM-related vocabulary), while groups AB had other 11 ATM jargon-based terms as part of the same test. The items that were not included here were designed as a different exercise during the same test.

However, another relevant class/test management factor might have been the 3 versus one row allocation of students.

The tables visibly show, however, despite the extremely high levels of proficiency, that precipitation seems to be the most difficult to assimilate efficiently, followed at a distance by snow, while air, wind, visibility and sky-cover seem to pose no problems at all.

3.2 Classic multiple-choice vocabulary exercise based on reading. This exercise is a simple reading designed for the main purpose of reinforcing essential vocabulary for cadet group studying aviation weather as the core of their respective curriculum. The worksheet comprised 6 multiple-choice grid response items, each made out of a stem (questions or partial sentences / incomplete statements) with 4 potential alternatives (answers) each, out of which 3 were distractors and one was the correct choice. In the case of items 3 and 4, the degree of difficulty was raised by the presence of strings (enumerations) of four terms and/or phrases, the invalidity of any resulting in the dismissal of the answer as incorrect, e.g.: *Stem 3. Choose the string which includes at least one term/phrase that does not refer to air movement. Answers: a) crosswind, tailwind, head wind, windshear; (…) c) downdraft, gradient wind, glaze, drift; d) (…),* where neither glaze nor drift are related to air movement and thus making answer c) into the right choice. Under item 5, further difficulty was added by the requirement to choose one correct match, to the purpose of reinforcing counter-intuitive abbreviations: *Stem 5. Find the wrong correspondence between the phenomenon and its METAR code. Answers: a) small hail/snow pellets/graupel = GS; b) hail = GR; c) small hail/snow pellets/graupel = GR; d) clear-sky precipitation / diamond dust = IC,* where the right answer was c).

No other linguistic context than the stems was provided in this classic multiple-choice vocabulary exercise. Previous lectures and topic-centered seminars had already preceded this reinforcement exercise. The content of the stems and grids basically remains upper intermediate-level in terms of difficulty: the stems are not basic, but are rather accessible while the specialist vocabulary in the answers is only slightly above the general military English vocabulary included in level 2 – intermediate STANAG 6001 descriptions; mention also need to be made that group D was studying aviation weather as a fundamental in its curriculum.

Learner answers turned out as described in Table 6, where each cadet was assigned a number and a group designator (in this case, D), to observe the privacy-related ethical concerns mentioned under section *2. Methodology*.

Table 6. Multiple-choice vocabulary exercise

	Std. code	*No. of errors / No. of answers*	*Details (based on answer key: 1c, 2d, 3c, 4b, 5c, 6a)*
1.	D1	1/6	• *1*, turbulence intensity*:* 1b, moderate (Correct: *1c, severe*)
2.	D2	0/6	–
3.	D3	2/6	• *1*, turbulence intensity*:* 1b, moderate (Correct: *1c, severe*) • *6*, runway conditions (water, clouds & precipitation)*:* undecisive marking (all four answers marked)
4.	D4	0/6	–
5.	D5	0/6	–
6.	D6	0/6	–
7.	D7	1/6	• *1*, turbulence intensity*:* 1b, moderate (Correct: *1c, severe*)
8.	D8	1/6	• *6*, runway conditions (water, clouds & precipitation)*:* undecisive marking (2 answers marked: *6a, pool of water* and 6b, standing water)
9.	D9	1/6	• *5*, METAR codes*:* 5d, clear-sky precipitation/diamond dust = IC (Correct: *5c, small hail/snow pellets/graupel = GR*)
10.	D10	1/6	• *4*, precipitation*:* undecisive marking (all four answers marked)

In group D, 60 answers received via the 10 worksheets included for assessment. The table demonstrates that certain unexpected difficulty associated with item number 1, focused on degrees of turbulence intensity, dominates the chart (3 wrong answers against

none under item 3, one under item 4, which would have been the most difficult for learners struggling with intermediate-level proficiency).

The stem was made out of an incomplete sentence to be filled in, reading: *1.* --------- *turbulence is described as disturbed air that may cause an aircraft to briefly drift in and out of control, while inside the aircraft walking becomes temporarily impossible.* This is a slightly rephrased definition of severe turbulence as compared to the two-fold general textbook definition provided in previous class materials, i.e., *Aircraft reaction: large and abrupt changes in altitude and/or attitude and, usually, large variations in indicated airspeed. The airplane may momentarily be out of control. Inside the aircraft: Occupants of the airplane will be forced violently against their seat belts. Food service and walking impossible*. The answers were made up of the modifiers in the noun phrases technically describing the classic 4-leveled turbulence intensity scale, from *light* to *extreme*.

Another improbable finding is the cadets' difficulty with item 6 (standing second just behind item 1 for having resulted in 2 wrong choices out of the total of 7). This was basically a classic, one-concept-per-answer find-the-best-synonym exercise focused on the terms describing different types and sizes of water build-ups (form standing water patches to running water and cloud build-ups) as runway conditions: *puddle, pool of water*, *standing water*, to which distractors describing cloud formations or running water descriptors had been added, namely *c) CB build-up* and *d) runoff*. Of course, the terms rely heavily on general/plain English vocabulary, but they do have rather specific, technical synonymies in the aeronautical jargon (for disambiguation, see the *Global Reporting Format (GRF) for Runway Surface Conditions* [9]) and may be rather counter-intuitive for Romanian learners (e.g., the synonymy between pool of water and puddle to describe standing water patches of similar shapes, surfaces and depths under the larger umbrella of the hypernym *standing water*). However, these details were one of the main focuses of the corresponding lecture and previous seminars. Thus, D8's hesitation is rather strongly motivated, while D3's is much less justifiable.

It is also relevant, perhaps, that just like the vocabulary around liquid and freezing precipitation, or the terms describing frozen water buildups in the case of the matching exercises, the semantic field of water (liquid runway build-ups, in this case) still holds one of the top ranking places.

Other than that, students in this group still performed very well, but with lower levels of performance than their AM peers in group C: out of the total of 60 answers, only about 11,66% were erroneous. Still, learners having made mistakes counted more than 60% of the group, while only 40% made no mistakes throughout. The lowest cadet performance was at approximately 66.66% correct answers. This may also pint out that this type of multiple-choice item grid is less learner-friendly than the classic matching exercises, even in spite of the fact that arithmetically, students are faced here with a higher chance of error since they need to manage a 4-answer, not an 11-answer choice per assignment.

4. CONCLUSIONS: SYNTHESIS AND FURTHER LINES OF INQUIRY

The quantitative error approach under section *3. Corpus analysis* readily and distinctly points out the areas of difficulty as far as the selective topics and related specialist vocabulary are concerned: with 22 out of the total number of 27 erroneous answers given by all the students tested on all the worksheets provided, i.e., a percentage of approximately 88,88% of all the mistakes made (which cand be broken into the 6 errors out of the 6 mistakes made by the 11 cadets in AM group C and 4 out of the total of 7 mistakes made by the 10 cadets in AM group D, plus the 85.71% or 12 out of the total number of 14 mistakes made by the 18 students in ATM groups A and B), the semantic

field of specialist terms related to water and precipitation (in the following order from precipitation METAR codes, leading with almost half of the errors committed (10 errors), followed by snow, ice and water buildups in a relatively equal amount (11 errors considered together) heads the chart.

While sky cover and visibility poses a virtually negligible threat (7.40% / 2 mistakes per 27), wind having a literal zero impact factor on error-making, but snow buildup representing a 14.81% percentage (4 out of 27 mistakes), and air movement made into an almost equally inconsequential threat (with 3 mistakes and 11.11% of the total), standing water as a runway condition is the only water and precipitation-related sub-topic falling in the same category as sky cover and visibility, with a 7.40% rate of error.

Thus, only the 22.72% of the total number of mistakes (i.e., 5 out of 22 mistakes) were fully unrelated to water (without even considering the presence of water under the form of clouds). Another area of concern may be pertaining to the classification of turbulence according to degrees of intensity, since all the errors under air movement concern the one corresponding item.

The numeric values included in the study are synthesized in Table 6 and results are detailed in Tables 7 and 8 for a better grasp on the various data around our topic and sub-topic list. Table 6 shows the proportions of right (94, 77.68%) and wrong (27, 22.31%) answers out of 121, accounting for general student performance throughout the four groups (A, B, C, and D) considered together. More detailed insights were not calculated as our objectives do not include student group performance assessment. However, the general proficiency demonstrated by the cadets involved will be relevant for future developments, as the same groups and students will be used (and their performance, relevant) for research concerning other activity types, as well as the qualitative interpretation of all results.

Table 6. Quantitative sum-up (general values)

General information	**Values**
total no of answers	121
total no of wrong answers	27
proportional value wrong answers	22.31%
total no of right answers	94
proportional value of correct answers	77.68%
total no of items	45

The last two sum-up tables (Table 7, Table 8) were generated by manually counting an introducing data into Microsoft Word documents and Microsoft Excel dedicated sheets. OpenAI's free ChatGPT website (https://chatgpt.com/) and an online free Percentage Calculator provided by calculator.net (https://www.calculator.net/percent-calculator.html) were also used for certain calculations (percentages, final line-up) and check-ups (i.e., result double-checking). While Table 7 synthesizes the final ranking of the topic and sub-topic related errors, based on the full calculations displayed under Table 8, as previously summarized at the beginning of this section, with values truncated to two decimal places without rounding them (as the same procedure has been applied consistently, for simplification, throughout). The general topics have been highlighted for better visibility and easy calculations.

Table 7 provides insights on the synthetic final ranking of topic difficulty based on the quantitative error analysis. All major topics and sub-topics are numerically present. It is visible how sky cover and air movement are largely surpassed by precipitation, with METAR codes, buildup and snow/ice at the top of the list and wind at the absolute bottom, with zero mistakes throughout.

Table 7. Quantitative sum-up (final ranking)

	Topics (highlighted) and sub-topics (regular)	Number of Errors	Proportional Value of Errors (%)
1.	Water and precipitation	22	81.48%
2.	METAR codes	9	33.33%
3.	Buildup	6	22.22%
4.	Snow/Ice	5	18.51%
5.	Runway Conditions	2	7.40%
5.	Sky Cover and visibility	2	7.40%
6.	Air Movement	3	11.11%
6.	Turbulence & Chop	3	11.11%
7.	Wind	0	0%

Table 8 provides a visible comparative take on items and the distribution of presence of topics and sub-topics among the items in all exercises, beside the number of errors and their proportional distribution around topics and sub-topics.

Table 8. Quantitative sum-up (details)

Topics (highlighted) and sub-topics (regular)	Items / total no. of items (45)		Errors / total no. of errors (27)	
water and precipitation	24	53.33%	22	81.48%
METAR codes	10	22.22%	9	33.33%
buildup	7	15.55%	6	22.22%
snow/ice	6	13.33%	5	18.51%
runway conditions	1	2.22%	2	7.40%
air movement	12	26.66%	3	11.11%
wind	10	22.22%	0	0.00%
turbulence & chop	2	4.44%	3	11.11%
sky cover and visibility	4	8.88%	2	7.40%
other*	5	11.11%	0	0.00%

Of course, we should also take into account the fact that out of the total of 45 items, approximately 24 (53.33%) were focused on or included vocabulary around water and precipitation (phases and their METAR codes, buildups of cloud, snow, ice, and standing water/runway conditions), about 12 (26.66%) around air movement (wind, turbulence and chop), 4 (8.88%) around sky cover and visibility conditions and 5 (11.11%) around other related vocabulary, e.g. drift. water buildups (snow, ice, cloud and standing water), about 3 (9.37%) around, and finally about again 10 (31.25%) around air movement (mostly wind, turbulence, chop).

But even so, the disproportion between topic and sub-topic related mistakes persists even when the number of items is considered (22 mistakes per 24 items revolving around water and precipitation, against 3 mistakes per 12 items about air movement, 2 errors per 4 items on sky cover and visibility etc.) and cannot be fully accounted for without a complementary quantitative analysis organized around less traditional activity types such as skill-focused tasks, where specialist terminology is provided with simulations of real-life communicative abilities and contexts and a qualitative sum-up based on both quantitative researches. The findings may seem disconcerting, as water may seem to attract error in an almost occult manner. Of course, nothing of the sort is to be really suspected: rational hypotheses may be mainly assumed to gravitate around the (user) unfriendliness of truncation and etymology, cultural and educational backgrounds, the presence or absence of real or realistically simulated linguistic and communicational contexts, learning habits such as memorizing and L1 interference. To be (or not to be) confirmed.

AKNOWLEDGMENT

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

This paper is supported by the Romanian MoND, under the The Ministry of National Defense Research & Development Sectorial Plan / 2023, project title: „Analiza erorilor în utilizarea limbii engleze folosite în aviație și comunicațiile aeronautice"; partners: „Henri Coandă" Air Force Academy and The Romanian Air Force Staff.

REFERENCES

[1] Å. Haukås, C. Bjørke and M. Dypedahl (Eds.), *Metacognition in Language Learning and Teaching*, Routledge, NY and London, 2018;

[2] X. Zhang, M. Guo, Metacognition and Second Language Learning, *Advances in Social Science, Education and Humanities Research*, vol. 428, pp.88-91, 2020.

[3] Y. In'nami, M. W.-L. Cheung , R. Koizumi and M. P. Wallace, Examining Second Language Listening and Metacognitive Awareness: A Meta-Analytic Structural Equation Modeling Approach, *Language Learning*, September, pp. 759–798, 2023;

[4] English as a foreign language learners' metacognitive experiences and writing development: Unraveling the process of learning EFL writing, *Learning and Individual Differences*, Volume 115, October, pp. 1-11, 2024. Available at https://pdf.sciencedirectassets.com/, accessed 04 Nov. 2024;

[5] M. Tavakoli, S. Hashemi Shahraki, M. Rezazadeh, The Relationship between Metacognitive Awareness and EFL Listening Performance: Focusing on IELTS Higher and Lower Scorers, *The Journal of Language Teaching and Learning*, 2, pp. 24-37, 2012;

[6] A.L. Wenden, Learner development in language learning, *Applied Linguistics*, 23 (1), pp. 32-55, 2002.

[7] L. Plonsky, *Advancing Quantitative Methods in Second Language Research*, Routledge, NY and London, 2015.

[8] A. Phakiti, P. De Costa, L. Plonsky and Sue Starfield (Eds.), Palgrave Macmillan, London, 2018.

[9]***Civil Aviation, Standards. Advisory Circular (AC) No. 300-019, Global Reporting Format (GRF) for Runway Surface Conditions, February 21, 2021. Available at https://tc.canada.ca/en/aviation/reference-centre/advisory-circulars , accessed 04 Nov. 2024.