Philologia LXVII
IANUARIE-IUNIE 2025

CZU:81’32
https://doi.org/10.52505/1857-4300.2025.1(324).06

Evaluarea abilității de scriere conform STANAG 6001: soluții digitale

Rodica MOȘNEGUȚU

Doctor în filologie
E-mail: rodicapinzaru@yahoo.com
ORCID: https://orcid.org/0009-0004-7916-3650
Academia Militară „Alexandru cel Bun” (Chișinău)

Evaluation of the Writing Skill in accordance with STANAG 6001: Digital Solutions

Abstract

The article examines the use of holistic and analytical assessments for English writing tests according to NATO STANAG 6001 standards. The author outlines the advantages and limitations of both methods, emphasizing the need for a review of digital tools used for evaluation. It suggests improvements to a digitalized evaluation form developed by the team at the Linguistic Center of “Alexandru cel Bun” Military Academy, which supports analytical performance measurement. This tool provides multiple benefits, such as more precise standardization of the evaluation process, reduction of correction time, and minimization of subjectivity in assessments. The contributions of the article highlight the importance of optimizing evaluation methods to ensure greater fairness and efficiency in assessing written language competencies.

Keywords: STANAG 6001 testing, holistic evaluation, analytical evaluation, test reliability, test validity, writing skill evaluation form.

Rezumat

Articolul analizează utilizarea evaluărilor holistică și analitică pentru testele de scriere în limba engleză conform standardelor NATO STANAG 6001. Autorul prezintă avantajele și limitările ambelor metode, subliniind necesitatea unei revizuiri a instrumentelor digitale utilizate pentru evaluare. Este propusă îmbunătățirea unui formular digitalizat, creat de echipa Centrului lingvistic al Academiei Militare „Alexandru cel Bun”, care sprijină evaluarea analitică a performanței lingvistice. Acest instrument oferă multiple beneficii, precum standardizarea mai exactă a procesului de evaluare, reducerea timpului necesar corectării și diminuarea subiectivității în aprecieri. Contribuțiile articolului evidențiază importanța optimizării metodelor de evaluare pentru a asigura corectitudinea și eficiența în evaluarea competențelor lingvistice scrise.

Cuvinte-cheie: testare STANAG 6001, evaluare holistică, evaluare analitică, fidelitate a testului, validitate a testului, formular de evaluare a abilității de scriere.

Evaluarea obiectivă a abilităților lingvistice într-o limbă străină reprezintă o provocare constantă pentru specialiștii implicați în testare. Rolul evaluărilor este de a le oferi candidaților cele mai bune condiții și cele mai potrivite subiecte, astfel încât aceștia să poată demonstra, la cel mai înalt nivel, competențele lingvistice. Pentru atingerea obiectivelor propuse, evaluatorii se raportează la un spectru complex de aspecte cercetate științific: fidelitate, validitate, conținutul testului, diferite abordări și criterii de testare etc. Acest articol examinează influența abordărilor *holistică* și *analitică* asupra obiectivității și validității evaluării abilității de scriere, conform standardelor STANAG 6001, precum și rolul unui instrument digitalizat în îmbunătățirea procesului respectiv.

Această lucrare abordează testarea lingvistică în conformitate cu criteriile de evaluare definite de descriptorii STANAG 6001 (ATrainP-5, 2016), care reprezintă standardul de referință pentru evaluarea competențelor lingvistice în cadrul NATO, elaborat în urma unei colaborări internaționale între experți din domeniul lingvisticii și al forțelor armate, având scopul de a asigura uniformitatea și comparabilitatea rezultatelor în contexte diverse. Structura sa derivă din necesitatea de a delimita în mod clar criteriile de apreciere specifice fiecărei abilități lingvistice: vorbire, audiere, citire și scriere.

Conform STANAG 6001, aceste abilități sunt evaluate pe cinci niveluri de performanță, cunoscute neoficial ca: nivelul 1 – de supraviețuire, nivelul 2 – funcțional, nivelul 3 – profesional, nivelul 4 – de expert și nivelul 5 – apropiat de cel al unui vorbitor nativ, cu abilități avansate de exprimare (*Ibidem*, p. A-1 – A-8). Fiecare nivel este descris prin indicatori preciși, care includ competențe lingvistice, gramaticale și de comunicare. Validitatea acestui cadru de evaluare a fost confirmată prin studii de corelare cu alte standarde internaționale, cum ar fi TOEIC (Tannenbaum, Baron 2010) și CEFR (Solak 2011), demonstrând aplicabilitatea descriptorilor în testarea lingvistică standardizată.

În plus, studiul realizat de J.D. Brown, J. McE. Davis și Chika Takahashi (2012) analizează relația dintre testele EIKEN, elaborate în Japonia, și TOEFL, recunoscut pe plan internațional, evidențiind posibilitatea utilizării acestor teste în paralel pentru admiterea la instituțiile de învățământ superior la nivel global. Totodată, cercetările din domeniul evaluării lingvistice subliniază importanța integrării unor metodologii riguroase pentru diferențierea și măsurarea competențelor lingvistice la fiecare nivel, asigurând o abordare echitabilă și obiectivă (Hughes 2003).

Pe parcursul procesului de evaluare, distincția între abilitățile de producere (vorbire și scriere) și cele de receptare (audiere și citire) rămâne semnificativă, din perspectiva gradului de obiectivitate implicat. Studiile de specialitate confirmă faptul că evaluarea abilităților de receptare are un caracter predominant obiectiv, prin utilizarea unor sarcini standardizate și itemi testați anterior (Bachman 1990, p. 74). Pe de altă parte, abilitățile de producere sunt semnificativ influențate de percepțiile subiective ale evaluatorilor, ceea

ce impune necesitatea folosirii unor metode riguroase de standardizare și a tehnologiei, astfel încât să fie redusă subiectivitatea.

Pentru a reduce influența subiectivității personalului de testare în procesul de evaluare, sunt implementate diverse metode de control, care includ introducerea unor criterii bine definite pentru măsurarea performanței candidaților. În acest context, se disting două abordări majore – *evaluarea holistică* și *evaluarea analitică* –, fiecare contribuind, în mod distinct, la asigurarea fidelității[1] (*reliability*) și validității[2] procesului de evaluare a probei scrise.

Abordarea holistică implică o apreciere generală a performanței de scriere a candidatului, bazată pe o impresie globală asupra răspunsului oferit (Hughes 2003, pp. 94-95). Cercetările indică faptul că obiectivitatea acestui tip de evaluare crește proporțional cu numărul de evaluatori implicați, patru fiind considerat un număr optim pentru asigurarea unei evaluări fidele și consistente. Această metodă oferă avantajul unei evaluări rapide și integratoare, însă presupune riscul unei variabilități interpretative în absența unui consens clar între evaluatori.

Un exemplu de evaluare holistică este scara de apreciere folosită de multe universități anglofone, care testează în baza unor criterii holistice capacitățile viitorilor studenți (*Ibidem*, p. 95). Scara de evaluare folosită de evaluatorii bine instruiți include următoarele niveluri: vorbitor nativ, aproape vorbitor nativ, evident mai mult decât corespunzător, posibil mai mult decât corespunzător, corespunzător pentru a-și efectua studiile la universitate, incert, clar necorespunzător, mult sub nivelul corespunzător. Totuși, acest tip de criterii de evaluare sunt concepute pentru un context anume și pot fi folosite doar în scopul pentru care au fost elaborate (Brown 2004, p. 242). Alte contexte vor necesita revederea scării respective și adaptarea la auditoriul-țintă și la obiectivele evaluării.

Atunci când se elaborează o scară de evaluare, aceasta trebuie să se bazeze pe cercetarea ordinii naturale de achiziție a elementelor de limbă, analizând etapele prin care un vorbitor non-nativ dobândește progresiv structuri gramaticale, elemente de vocabular și competențe de comunicare. În absența unei astfel de cercetări, testele riscă să nu măsoare în mod corect și echitabil abilitățile candidaților, afectând astfel fidelitatea și validitatea rezultatelor obținute. Stephen Krashen (1982, p. 12), un nume de referință în studierea proceselor de achiziție a unei limbi, susține că dobândirea limbajului urmează o secvență predictibilă, indiferent de limba-țintă, idee detaliată în lucrările sale.

---

[1] Obținerea unor rezultate similare fie în urma administrării repetate a testului aceloraşi persoane (testare/testare repetată), fie a corectării sau notării acestuia de către examinatori diferiți ori de către același examinator, după scurgerea unui interval de timp (notare/notare repetată) (Ursa, Marta 2015, pp. 151-160).

[2] Validitatea se referă la capacitatea unui test de a măsura ceea ce trebuie să măsoare (Hughes 2003, p. 26).

În contextul evaluării practice, apare adesea dilema plasării candidaților la un anumit nivel de competență, în special atunci când aceștia îndeplinesc parțial criteriile corespunzătoare pentru două niveluri proxime. A. Hughes (2003, p. 100) subliniază că această decizie trebuie să fie influențată de scopul evaluării. Dacă testarea este realizată în contextul selecției pentru ocuparea unei funcții sau a unui rol cu responsabilități importante, evaluatorii ar trebui să încline spre situarea candidatului la nivelul inferior, deoarece miza deciziei este semnificativă și necesită precauție. În schimb, în cazul unei evaluări formative, care vizează monitorizarea progresului educațional, anumite deficiențe pot fi compensate prin abilități demonstrate la un nivel superior, asigurând astfel o evaluare mai flexibilă și centrată pe dezvoltare.

Abordarea analitică implică evaluarea unei performanțe lingvistice prin analiza mai multor dimensiuni ale limbii, cum ar fi corectitudinea gramaticală, organizarea textuală, coerența și ortografia. Prin acordarea unui scor distinct pentru fiecare criteriu și calcularea mediei rezultatelor obținute, această metodă permite o apreciere mai detaliată și cuprinzătoare (Brown 2004, p. 243).

Printre avantajele abordării analitice se numără următoarele:

- acordarea unei atenții sporite unor criterii care, în alte condiții, ar putea fi ignorate;
- reducerea problemelor legate de o dezvoltare inegală a sub-abilităților candidaților;
- creșterea fidelității evaluării prin obligativitatea de a acorda un scor pentru fiecare criteriu specific.

Totuși, există riscul ca evaluatorii să nu poată analiza fiecare criteriu în mod complet independent. Cu toate acestea, includerea mai multor dimensiuni evaluative conferă o fidelitate mai ridicată rezultatelor finale (Hughes 2003 pp. 102-103). De asemenea, diverse aspecte evaluate pot avea ponderi diferite: de exemplu, corectitudinea ortografică poate fi considerată mai puțin importantă decât acuratețea gramaticală.

Principalul dezavantaj al metodei analitice constă în faptul că implică un consum considerabil de timp pentru evaluatori, inclusiv pentru cei cu experiență. De asemenea, concentrarea exclusivă pe evaluarea fragmentată a diferitelor criterii poate diminua capacitatea evaluatorului de a aprecia răspunsul scris în ansamblu. În astfel de situații, combinarea cu o evaluare holistică devine necesară pentru a asigura un echilibru între fidelitate și validitate. În absența unei perspective generale, metoda analitică riscă să ofere rezultate precise, dar care nu reflectă în mod adecvat calitatea reală a performanței. D.H. Brown (2004, p. 426) sugerează utilizarea evaluării holistice pentru testele de evaluare generală, în timp ce evaluarea analitică este mai potrivită pentru testele formative, întrucât oferă informații detaliate despre progresul studenților.

Evaluarea abilității de scriere conform standardelor STANAG 6001 la Centrul lingvistic al Academiei Militare „Alexandru cel Bun” a cunoscut două etape distincte.

Inițial, s-a aplicat metoda analitică de apreciere, iar ulterior s-a adoptat evaluarea holistică, utilizată în prezent. Studiul de față examinează soluții electronice pentru optimizarea procesului de evaluare analitică, având scopul de a reduce timpul și efortul evaluatorilor și de a spori obiectivitatea rezultatelor. Prin integrarea acestor metode, se urmărește creșterea fidelității și validității evaluărilor lingvistice.

Testarea abilității de scriere a evidențiat două deficiențe majore ale metodei analitice în momentul aplicării. Primul și cel mai evident neajuns a fost lipsa digitalizării formularelor de evaluare, în ciuda utilizării unor expresii standardizate pentru descrierea performanței candidaților. Acest aspect a determinat încetinirea considerabilă a procesului de apreciere. Evaluarea fiecărui criteriu – sarcini, organizare, vocabular, gramatică și sintaxă – implica atribuirea unor formulări apreciative corespunzătoare nivelurilor 0-3. Reluarea constantă a acelorași formulări în procesul de corectare sublinia necesitatea dezvoltării unui program informatic, ce ar fi integrat și categorisit structurile lingvistice evaluative. Un astfel de sistem ar fi facilitat selecția automată a descriptorilor relevanți, oferind o caracterizare complexă a fiecărei lucrări, reducând semnificativ timpul alocat fiecărui candidat.

Soluția de digitalizare a fost identificată în cadrul aplicației Microsoft InfoPath, inclusă în suita Microsoft Office 2003-2013, care permite, printre altele, crearea de câmpuri cu liste derulante (eng. *drop-down lists*), facilitând automatizarea procesului de evaluare. În acest sens, am inițiat procesul de colectare și organizare a expresiilor standard utilizate pentru aprecierea lucrărilor scrise, clasificându-le conform principiului de excludere mutuală (vezi Fig. 1).

Date Reviewer PIN # 0

Candidate's name

Level/ Task	Task Completion	Organization	Vocabulary	Grammar and Syntax	Level Awarded
Level 1	Notable content omi	Too little langague	Basic and high frequen	Very many errors i	
Task One	content adequate;	lack of sentence co	chunks lifted from the p	tenses, articles, pr	
	appropriate text typ	composition of texi	appropriate to the cont	some word order e	
	length requirements	introduction missin	very many errors;	the errors block th	2
	can		which block natural flo		
	state facts;				
	0	0	0	0	
Level 2	Notable content omi	Too little langague	Basic and high frequen	Very many errors i	
Task Two	content adequate;	lack of sentence co	expressed content in h	tenses, articles, pr	
	appropriate text typ	composition of texi	appropriate to the cont	good control of sin	3
	length requirements	introduction missin	very many errors;	the errors block th	
	can		which block natural flo		
	state facts;				
	0	0	0	0	
				Final result	2

Fig. 1. Formularul de evaluare a abilității de scriere la limba engleză

Mai exact, expresiile care descriu același criteriu, dar reflectă o progresie graduală a nivelului de performanță, au fost incluse într-o singură listă derulantă. De exemplu, în rubrica VOCABULAR, au fost definite următoarele categorii (Fig. 2): *pertinența în context, caracterul iterativ al vocabularului, tipologia vocabularului utilizat, frecvența erorilor lexicale și impactul acestora asupra înțelegerii textului*.

Una dintre listele derulante (Fig. 3) permite selecția dintre următoarele opțiuni:
– corespunde contextului;
– alegere și utilizare eficientă a cuvintelor și expresiilor idiomatice;
– utilizare necorespunzătoare a vocabularului;
– vocabular din afara contextului;
– uneori are loc utilizarea necorespunzătoare a vocabularului.

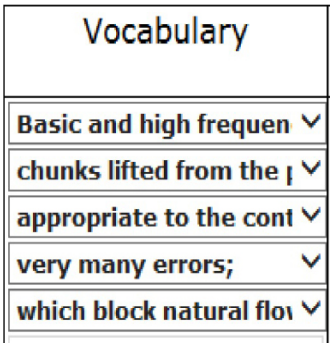


Fig. 2. Categorii de evaluare a vocabularului

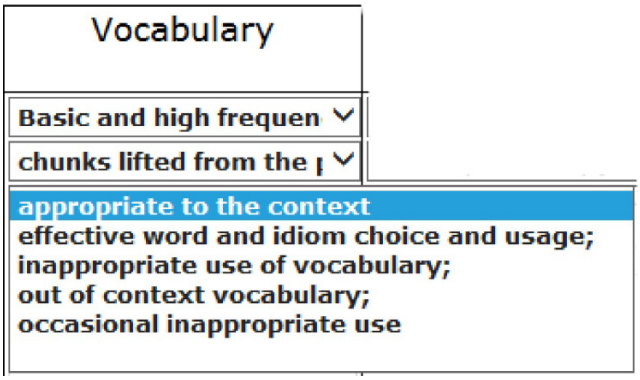


Fig. 3. Lista derulantă cu referire la pertinența vocabularului în context

Formulările de mai sus descriu spectrul de performanță lexicală în raport cu criteriul *pertinența în context*, utilizat în lucrarea analizată. Acestea corespund diferitelor criterii din descriptorii STANAG 6001, astfel încât selecția unei anumite opțiuni indică nivelul la care candidatul se încadrează pentru acest criteriu.

În anumite cazuri, au fost utilizate două liste derulante pentru a exprima o singură apreciere evaluativă, deoarece unele formulări conțineau două dimensiuni distincte ale evaluării, care se combinau diferit în funcție de lucrarea analizată. De exemplu, pentru a clasifica numărul erorilor și impactul acestora asupra comprehensiunii textului, au fost create două liste separate.

Prima listă oferă o evaluare cantitativă a erorilor și include opțiunile de apreciere: *foarte multe erori, puține erori ortografice, erori ocazionale, puține erori nesemnificative* (Fig. 4). A doua listă (Fig. 5) se concentrează pe influența erorilor asupra comprehensiunii textului și conține formulări precum: *împiedică lectura fluentă și comprehensiunea, distorsionează uneori semnificația, alegerea și utilizarea vocabularului fac dificilă uneori înțelegerea sensului, problemele de*

corespundere afectează frecvent înțelegerea, nu perturbă semnificația, vocabularul asigură claritatea mesajului.

Unele dintre opțiunile din cea de-a doua listă pot funcționa independent, în timp ce altele sunt utilizate în combinație cu formulările din prima listă, completând fragmentul evaluativ printr-o relație cauzală (ex. *care*…).

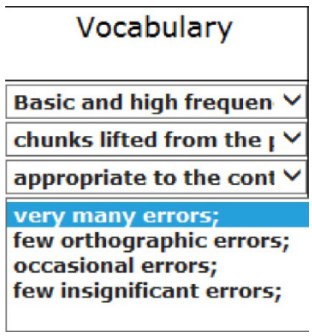


Fig. 4. Incidența erorilor de vocabular

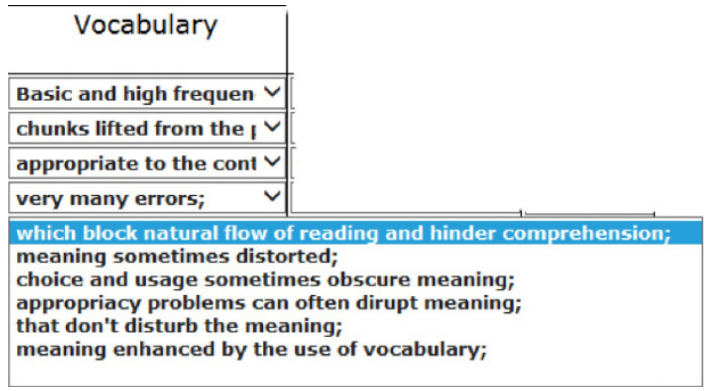


Fig. 5. Influența vocabularului utilizat în lucrare asupra înțelegerii conținutului

Toate capitolele de evaluare includ câteva categorii care permit o analiză detaliată a fiecărui aspect evaluat: SARCINI – puncte tratate din sarcină, relevanța conținutului, corespunderea tipului de text, lungimea, capacitatea de a produce anumite tipuri de text; ORGANIZARE – coerență, coeziune, organizare, structură; GRAMATICĂ ȘI SINTAXĂ – numărul erorilor, tipul erorilor gramaticale, tipul erorilor sintactice, impactul erorilor asupra înțelegerii.

În funcție de criteriile de apreciere selectate în formularul electronic, examinatorul atribuie nivelul STANAG 6001 (între 0 și 3) corespunzător aspectului evaluat pentru fiecare item. Apoi, se calculează un nivel general pentru fiecare dintre itemii din lucrare, iar, în final, nivelurile obținute sunt analizate și candidatului îi este oferit rezultatul final pentru lucrarea scrisă.

Automatizarea parțială a procesului de evaluare a probei scrise în cadrul testului STANAG 6001 a condus la dezvoltarea unui program destinat eficientizării verificării și acordării nivelurilor lingvistice. Acesta integrează formulări frecvent utilizate în aprecierea performanței candidaților și evidențiază structurile lingvistice relevante pentru evaluare. Instrumentul facilitează selecția opțiunilor care descriu cel mai fidel aspectele analizate, promovează o abordare sistematică și sugerează nivelul de competență al candidatului, contribuind astfel la creșterea obiectivității evaluării.

Adoptarea formularului digitalizat dezvoltat de echipa Centrului Lingvistic al Academiei Militare „Alexandru cel Bun” reprezintă o inovație importantă în procesul de evaluare analitică. Standardizarea sporită a criteriilor reduce variabilitatea interpretărilor, eficientizează timpul alocat corectării și asigură o aplicare uniformă.

În plus, diminuarea subiectivității în evaluare consolidează validitatea și fidelitatea rezultatelor, candidații beneficiind de o apreciere corectă și echitabilă.

Pe termen lung, utilizarea acestui formular sprijină nu doar procesul de evaluare, ci și formarea continuă a evaluatorilor. Aplicarea constantă a instrumentului contribuie la calibrarea competențelor și la implementarea unor practici standardizate, ce pot fi folosite în diverse contexte profesionale.

În vederea optimizării, formularul ar putea fi îmbunătățit prin integrarea expresiilor evaluative derivate direct din descriptorii STANAG 6001, sporind astfel fidelitatea aprecierilor. Asocierea fiecărei formulări cu nivelul corespunzător de competență ar adăuga o dimensiune numerică evaluării, iar listele derulante ar putea include opțiuni prestabilite pentru nivelurile 0, 0+, 1, 1+, 2, 2+ și 3. Totodată, utilizarea formulelor pentru calcularea automată a nivelului final ar spori obiectivitatea și acuratețea procesului.

Aplicația Microsoft InfoPath oferă oportunități pentru interconectarea câmpurilor și calculul automat al nivelurilor, contribuind la dezvoltarea unei platforme integrate. Validarea acestei ipoteze necesită cercetări suplimentare, care ar putea conduce la crearea unui formular de evaluare analitică aliniat la standardele STANAG 6001.

În ciuda avantajelor metodei analitice, literatura de specialitate subliniază că suma părților nu reflectă întotdeauna ansamblul. Pentru o evaluare validă și reprezentativă, aceasta trebuie completată cu o lectură holistică, o abordare globală, ce contribuie la eliminarea fragmentării excesive și la obținerea unei aprecieri echilibrate.

Implicarea mai multor evaluatori în procesul de evaluare holistică sporește fidelitatea rezultatelor și reduce subiectivitatea individuală, oferind interpretări mai nuanțate. În plus, metoda analitică rămâne esențială pentru formarea evaluatorilor și calibrarea periodică a standardelor, sprijinind atât selecția profesională, cât și monitorizarea progresului educațional.

Astfel, integrarea evaluării analitice și holistice oferă un echilibru necesar între părți și ansamblu, asigurând validitatea rezultatelor și adaptabilitatea acestora la diverse scopuri. Combinația celor două metode reflectă complexitatea competențelor lingvistice ale candidaților și contribuie la dezvoltarea unor sisteme evaluative echitabile și eficiente.

Referințe bibliografice:

ATrainP-5. NATO STANAG 6001: Language Proficiency Levels. Edition A Version 2. North Atlantic Treaty Organization: Standardization Office, 2016, p. 32. Disponibil: https://natobilc.org/wp-content/uploads/2024/11/ATrainP-5-EDA-V2-E.pdf [accesat 2025-03-28].

BACHMAN, Lyle F. *Fundamental Considerations in Language Testing.* Oxford: Oxford University Press, 1990. ISBN 0 19 437003 8. Disponibil: https://www.academia.edu/28794667/Fundamental_Considerations_in_Language_Testing [accesat 2025-03-28].

BACHMAN, Lyle F. and PALMER, Adrian S. *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press, 1996. ISBN 0 19 437148 4. Disponibil: https://octovany.wordpress.com/wp-content/uploads/2013/12/language-testing-in-practice-bachman-palmer.pdf [accesat 2025-04-01].

BROWN, H. Douglas. *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education Inc., 2004. ISBN 0-13-098834-0. Disponibil: https://evelintampubolon.wordpress.com/wp-content/uploads/2016/09/h-_douglas_brown_-_language_assessment.pdf [accesat 2025-03-31].

BROWN, James Dean; DAVIS, John McE. and TAKAHASHI, Chika. *Upper-level EIKEN Examinations: Linking, Validating, and Predicting TOEFL® iBT Scores at Advanced Proficiency EIKEN Levels.* University of Hawaii at Mānoa: Eiken Foundation of Japan, 2012. Disponibil: https://www.eiken.or.jp/eiken/group/result/pdf/eiken-toeflibt-report.pdf [accesat 2025-04-02].

HUGHES, Arthur. *Testing for Language Teachers: English language teaching for adults and children: ELT general.* Cambridge: Cambridge University Press, 2003. ISBN 978-0-521-48495-4.

KRASHEN, Stephen D. *Principles and Practice in Second Language Acquisition.* University of Southern California: Pergamon Press Inc., 1982. ISBN 0-08-028628-3. Disponibil: https://www.sdkrashen.com/content/books/principles_and_practice.pdf [accesat 2025-03-28].

SOLAK, Ekrem. Historical Development of NATO STANAG 6001 Language Standards and Common European Framework (CEF) and The Comparison of Their Current Status. In: Gümüşhane Üniversitesi, Sosyal Bilimler Elektronik Dergisi, 2011. Disponibil: https://files.eric.ed.gov/fulltext/ED553404.pdf [accesat 2025-04-01].

TANNENBAUM, Richard J. and BARON, Patricia A. *Mapping TOEIC Test Scores to the STANAG 6001 Language Proficiency Levels*. Princeton, New Jersey: Educational Testing Service, 2010. Disponibil: https://www.ets.org/Media/Research/pdf/RM-10-11.pdf [accesat 2025-04-01].

URSA, Ovidiu și MARTA, Monica Mihaela. Caracteristici şi limitări ale testelor de evaluare a competenţei lingvistice. În: Buletin Ştiinţific, Fascicula filologie, Seria A, vol. XXIV, 2015, pp. 151-160. ISSN 1583-1264. Disponibil: https://www.researchgate.net/publication/286935904_Caracteristici_si_limitari_ale_testelor_de_evaluare_a_competentei_lingvistice_Characteristics_and_limitations_of_language_proficiency_tests [accesat 2019-10-02].

Primit: 17.02.2025
Acceptat: 01.05.2025