

УДК811. 111(07)

## КОМП'ЮТЕРНЕ ТЕСТУВАННЯ У НАВЧАННІ ІНОЗЕМНИХ МОВ: АНОТОВАНА БІБЛІОГРАФІЯ

Плотніков Є. О., Ларіна Т. В.

plotnikov@ndu.edu.ua; larinatanya300870@gmail.com

Ніжинський державний університет імені Миколи Гоголя

Дата надходження 20.03.2019. Дата прийняття до друку 20.05.2019.

У статті зроблено спробу короткого аналізу розвитку галузі використання комп'ютерних технологій для тестування іншомовних комунікативних компетентностей. Подано анотований огляд публікацій зазначеної тематики у провідних наукових виданнях з методики навчання іноземних мов за період з 2010 до 2019 рр.

**Ключові слова:** комп'ютерне тестування, дослідження, науковий журнал.

**Плотников Е. А., Ларина Т. В. Нежинский государственный университет имени Николая Гоголя**  
**Компьютерное тестирование в обучении иностранным языкам: аннотированная библиография**

В статье предпринята попытка краткого анализа развития исследований в области применения компьютерных технологий тестирования иноязычных коммуникативных компетентностей. Предоставлен аннотированный обзор публикаций указанной тематики в ведущих научных изданиях по методике обучения иностранным языкам за период с 2010 по 2019 гг.

**Ключевые слова:** компьютерное тестирование, исследования, научный журнал.

**Plotnikov Y., Larina T. Nizhyn Mykola Gogol State University**

**Computer-based Language Testing in Scientific Periodical Publications: Annotated Bibliography**

This paper aims to investigate the development of the computer-based testing research field in the recent years. The article systematically reviews the key studies conducted and published from 2010 to 2019 to reveal current trends and to show best research practices. All the analyzed publications represent the leading scientific journals in the field of language education.

**Key words:** computer-based testing (CBT), research, scientific journal.

Комп'ютерне (або комп'ютеризоване) тестування іншомовної комунікативної компетентності – галузь методики навчання іноземних мов, яка динамічно розвивається, постійно висуваючи нові дослідницькі проблеми та способи їх вирішення. У статті подано огляд публікацій у провідних наукових періодичних виданнях, присвячених цій темі. З топ 10% категорії “*Language & Linguistics*” (2018) розглянуто журнали, публікації у яких виключно або переважно пов'язані з тематикою навчання іноземних мов й оволодіння ними. Крім того, у категорії “*Language & Linguistics*” виділимо ще дві групи журналів, що варті окремої уваги у контексті теми огляду. Перша з них – це журнали, присвячені мовному оцінюванню та тестуванню: *Language Testing* (50 місце, видавець – *SAGE Publications*), *Assessing Writing* (69, *Elsevier Ltd.*), *Language Assessment Quarterly* (146, *Lawrence Erlbaum Associates Inc.*). Друга група – це журнали, сфокусовані на комп'ютерних технологіях навчання іноземних мов, з-поміж яких: *Computer Assisted Language Learning* (23, *Taylor & Francis*), *Language Learning and Technology* (29, *University of Hawaii Press*), *ReCALL* (35, *Cambridge University Press*), *Teaching English with Technology* (226, *IATEFL Poland. Computer Special Interest Group*), *JALT CALL Journal* (258, *The JALT CALL SIG*), *CALL-EJ* (297, *Asia-Pacific Association for Computer-Assisted Language Learning*), *International Journal of Computer-Assisted Language Learning and Teaching* (353, *IGI Global Publishing*) тощо. Журнали обох груп додано до нашого переліку незалежно від їхнього місця у загальному рейтингу. Загалом до бібліографії увійшло 65 статей з 23 журналів. Період публікацій – 2010-2019 рр.

Зазначимо, що перелік не є вичерпним, оскільки поняття контролю й оцінювання більшою або меншою мірою згадуються у багатьох публікаціях. У нашому огляді ми виокремили лише статті, безпосередньо зосереджені на проблемі використання комп'ютерних технологій для тестування мовних / мовленнєвих комунікативних компетентностей або на таких, що приділяють їй значну увагу.

#### 2010

Baturay, M. H., & Daloplu, A. (2010). E-portfolio assessment in an online English language course. *Computer Assisted Language Learning*, 23(5), 413–428. <https://doi.org/10.1080/09588221.2010.520671>

Дослідження сконцентроване на проблемі використання електронного портфоліо як інструменту оцінювання іншомовного писемного мовлення і має на меті з'ясування ступеня його впливу на розвиток компетентності студентів у письмі. Експеримент продемонстрував помітний приріст рівня письмових умінь і виявив загальне позитивне ставлення студентів до ведення електронного портфоліо.

Franco, H., Bratt, H., Rossier, R., Gadde, V. R., Shriberg, E., Abrash, V., & Precoda, K. (2010). Eduspeak®: A speech recognition and pronunciation scoring toolkit for computer-aided language learning applications. *Language Testing*, 27(3), 401–418. <https://doi.org/10.1177/0265532210364408>

Система *EduSpeak®* – це інструментарій для розроблення програмного забезпечення, що дозволяє розробникам програмного забезпечення для навчання іноземних мов використовувати сучасну технологію розпізнавання мови й вимови. Автоматичне оцінювання вимови дає змогу комп'ютеру надавати зворотний зв'язок щодо загальної якості вимови й указувати на конкретні проблеми. У статті розглянуто цей підхід до оцінювання іншомовної вимови. Результати доводять, що класифікаційна похибка таких систем лише трохи перевищує показники середнього парного розходження для людських експертів.

Chapelle, C. A., Chung, Y. R., Hegelheimer, V., Pendar, N., & Xu, J. (2010). Towards a computer-delivered test of productive grammatical ability. *Language Testing*, 27(4), 443–469. <https://doi.org/10.1177/0265532210367633>

Пілотне дослідження комп'ютеризованих тестових завдань для перевірки англійських продуктивних граматичних навичок. Завдання розглянуто з точки зору 1) надійності, 2) зв'язку між складністю завдань й етапами розроблення, 3) кореляції з іншими тестами, 4) передбачуваності результатів тесту щодо рівнів володіння мовою. Результати підтвердили потенціал тестових завдань для оцінювання продуктивних граматичних навичок за допомогою комп'ютерних технологій.

#### 2011

Xi, X., & Mollaun, P. (2011). Using raters from India to score a large-scale speaking test. *Language Learning*, 61(4), 1222–1255. <https://doi.org/10.1111/j.1467-9922.2011.00667.x>

У фокусі дослідження оцінювання усної частини онлайн-тесту *Test of English as a Foreign Language (TOEFL iBT)* людьми, рідною мовою яких не є англійська. Запропоновано й перевірено ефективність спеціального курсу підготовки таких тестувальників. Розглянуто, в який спосіб після проходження навчального курсу індійські тестувальники змогли послідовно та якісно оцінювати носіїв індійських й інших мов під час складання *TOEFL iBT*. Продемонстровано, що якість роботи таких тестувальників відповідала загальним вимогам тесту та не відрізнялася від роботи фахівців – носіїв мови.

Blake, R. J. (2011). Current trends in online language learning. *Annual Review of Applied Linguistics*, 31, 19–35. <https://doi.org/DOI: 10.1017/S026719051100002X>

Розглядаючи основні напрямки надання й отримання мовної освіти онлайн (такі як синхронна й асинхронна онлайн комунікація, онлайн вправи, гейміфікація тощо), автор побіжно згадує деякі аспекти комп'ютеризованого тестування й оцінювання в межах цих напрямків.

Spinner, P. (2011). Second language assessment and morphosyntactic development. *Studies in Second Language Acquisition*, 33(4), 529–561. <https://doi.org/DOI: 10.1017/S0272263111000301>

Розглянуто особливості різних процедур оцінювання рівня морфо-синтаксичного розвитку людини. Деякі з процедур здійснюються з використанням комп'ютерних технологій. Так, подано загальний опис комп'ютеризованих усних тестів і швидкого вимірювання комунікативного профілю за Пінеманом.

## 2012

Stoynoff, S. (2012). Research agenda: Priorities for future research in second language assessment. *Language Teaching*, 45(2), 234–249. <https://doi.org/DOI: 10.1017/S026144481100053X>

Зроблено спробу аналізу ключових аспектів оцінювання й тестування іншомовної комунікативної компетентності. Окреслено чотири пріоритетні напрямки досліджень у цій галузі, одним з яких зазначено використання технологій у тестуванні. Сформульовано декілька дослідницьких завдань, спрямованих на перевірку впливу персональних особливостей тестованих (стать, соціальний статус, мотивація тощо) на результати тестування або технічних особливостей тесту на процес його виконання.

Hung, S. T. A. (2012). A washback study on e-portfolio assessment in an English as a Foreign Language teacher preparation program. *Computer Assisted Language Learning*, 25(1), 21–36. <https://doi.org/10.1080/09588221.2010.551756>

В основі дослідження – проект запровадження електронного портфоліо як методу альтернативного оцінювання в курсі підготовки майбутніх учителів. Виявлено як позитивний, так і негативний зворотний вплив використання електронних портфоліо на процес навчання. Так, використання портфоліо сприяло створенню навчальної спільноти, засвоєнню навчальної інформації, розвитку професійних умінь і критичного мислення. При цьому у деяких студентів спостерігалось зростання тривожності й схильність до заперечення використання комп'ютерних технологій у навчанні.

Teo, A. (2012). Promoting EFL students' inferential reading skills through computerized dynamic assessment. *Language Learning & Technology*, 16(3), 10–20.

За допомогою програми *Viewlet Quiz 3* розроблено засіб комп'ютеризованого динамічного оцінювання з метою поєднання медіації й оцінювання під час навчання іншомовного читання. Детально описано засіб разом із заходами його експериментальної перевірки та рекомендаціями щодо використання.

Goldin, I. M., & Ashley, K. D. (2012). Eliciting formative assessment in peer review. *Journal of Writing Research*, 4(2), 203–237. <https://doi.org/10.17239/jowr-2012.04.02.5>

Зроблено спробу порівняння ефективності різних типів виставлення рейтингів студентами у межах комп'ютеризованої системи взаємного оцінювання. Кожен з типів (проблемно й тематично орієнтовані) спирається на різні набори критеріїв для оцінювання письма. Експериментальним шляхом встановлено, що студенти розрізняють різні критерії, й обидва типи рейтингів корелюють з оцінками викладача. При цьому проблемно орієнтовані рейтинги видаються більш корисними й менш переобтяженими надмірною інформацією.

Odo, D. (2012). Computer familiarity and test performance on a computer-based cloze ESL reading assessment. *Teaching English with Technology*, 12 (3), 18-35. Retrieved from <http://www.tewtjournal.org/?wpdmact=process&did=MzM3LmhvdGxpbms>

Передбачено усунення прогалін у попередніх дослідженнях шляхом вивчення взаємозв'язку між рівнем умінь володіння комп'ютером і результатами комп'ютерного тестування іншомовного читання (інтегративного, а не дискретного). Досліджено компетентність у читанні онлайн учнів середньої школи одного з міст на заході Канади. Учні відповіли на запитання опитувальника про їхній рівень умінь володіння комп'ютером, а потім виконали онлайн-тест у заповненні пропусків. На відміну від більшості результатів інших досліджень, результати показали, що рівень умінь володіння комп'ютером справді становлять невелику, але значну частину варіативності результатів комп'ютерного тестування.

Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me?: Individual self-assessment by skill and level of language instruction. *System, 40(1)*, 144–160. <https://doi.org/https://doi.org/10.1016/j.system.2012.01.003>

Розглянуто самооцінювання студентів на трьох рівнях (початковому, середньому й високому) володіння мовою й запропоновано матеріали для перевірки взаємозв'язку між інструментами самооцінювання й мовним онлайн-тестом. Результати дослідження можуть бути корисними для виявлення здатності просунутих студентів оцінювати себе за наявності конкретних критеріїв.

### 2013

Barkaoui, K., Brooks, L., Swain, M., & Lapkin, S. (2013). Test-takers' strategic behaviors in independent and integrated speaking tasks. *Applied Linguistics, 34(3)*, 304–324. Retrieved from <http://dx.doi.org/10.1093/applin/ams046>

Розглянуто стратегії поведінки під час складання усної частини комп'ютерного тесту TOEFL і BT, а також зв'язок між поведінкою й отриманими результатами тестування. З'ясовано, що тестовані використовують більше стратегій під час роботи над інтегрованими завданнями (одночасно поєднують щонайменше два види мовленнєвої діяльності, наприклад, читання й говоріння). Не виявлено помітного зв'язку між використаними стратегіями й отриманими результатами. Акцентовано, що складання усної частини онлайн-тесту висуває особливі вимоги до говоріння та стратегій, яких воно потребує.

Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during Computerized Dynamic Assessment (C-DA). *Language Teaching Research, 17(3)*, 323–342. <https://doi.org/10.1177/1362168813482935>

Використано принципи динамічного оцінювання в онлайн-тестах для перевірки рівня розуміння сприйнятої інформації під час англomовних аудіювання й читання. Тести містять спеціальні елементи, що допомагають визначити, якою мірою під час тесту підтримується розвиток учня. Ці процеси перевіряються у трьох напрямках оцінювання: фактичному (безпосередня діяльність), опосередкованому (реакція на отриману допомогу) й оцінці навчального потенціалу (обсяг зусиль, які необхідно вкласти у майбутню навчальну діяльність).

Poehner, M. E., & Van Compernelle, R. A. (2013). L2 development around tests: Learner response processes and dynamic assessment. *IRAL - International Review of Applied Linguistics in Language Teaching, 51(4)*, 353–377. <https://doi.org/10.1515/iral-2013-0015>

Досліджено реакції й відповіді студентів під час виконання комп'ютеризованого динамічного оцінювання вмінь іншомовного читання. Запропоновано кейс-аналіз студентів бакалаврату, які вивчають французьку мову як іноземну. Стверджується, що під час динамічного оцінювання реакції студентів екстерналізуються завдяки діалогу з екзаменатором, який не тільки документує ці процеси, але й бере в них активну участь. Саме екзаменатор регулює орієнтацію тестованих на виконання завдання та створює можливості для навчання під час оцінювання.

Wang, Y. J., Shang, H. F., & Briody, P. (2013). Exploring the impact of using automated writing evaluation in English as a foreign language university students' writing. *Computer Assisted Language Learning, 26(3)*, 234–257. <https://doi.org/10.1080/09588221.2012.655300>

Виявлено вплив використання автоматизованої системи оцінювання письма (*Automated Writing Evaluation*) на розвиток компетентності у письмі студентів першого курсу мовного університету. За допомогою кількісних та якісних методів дослідження з'ясовано, що вміння студентів експериментальної групи, які використовували зазначену систему, розвивалися швидше, особливо правильність письма й розвиток навчальної автономії.

Jamieson, J., Wang, L., & Church, J. (2013). In-house or commercial speaking tests: Evaluating strengths for EAP placement. *Journal of English for Academic Purposes, 12(4)*, 288–298. <https://doi.org/10.1016/j.jeap.2013.09.003>

Описано вибір тесту для використання в курсі інтенсивного навчання англійської мови як іноземної з-поміж таких тестів, як: *PIE Speaking* (тест внутрішньої розробки) і *Versant*

*English* (стандартизований, комп'ютеризований тест; розробник – *Pearson Knowledge Technologies*). Тести порівняно за такими параметрами: відповідність навчальній програмі, статистичний розподіл і практичність. Обидва тести не повністю відповідали програмі. Аналіз тестів дав змогу дійти низки висновків: *PIE Speaking* за своїми показниками більш придатний для тестування студентів з низьким і високим рівнем володіння мовою, *Versant English* – для студентів з середнім рівнем.

DeLuca, C., Cheng, L., Fox, J., Doe, C., & Li, M. (2013). Putting testing researchers to the test: An exploratory study on the TOEFL iBT. *System, 41*(3), 663–676. <https://doi.org/https://doi.org/10.1016/j.system.2013.07.010>

Проаналізовано потенційно проблемні питання щодо комп'ютеризованого тесту TOEFL iBT, пов'язані з конструктивно залежними і нерелевантними параметрами варіації. Чотири фахівці з мовного тестування повідомили про свій досвід виконання тесту TOEFL iBT у двох фоку-групах. Їхній досвід опрацьовано за допомогою стандартного тематичного аналізу й згруповано на основі вимірювання варіації. Результати дослідження мають цінність щодо вимірювання англійської мови академічного спілкування і можуть бути використані для розв'язання практичних проблем, пов'язаних з розробленням і використанням тесту TOEFL iBT.

Newhouse, C. P., & Cooper, M. (2013). Computer-based oral exams in Italian language studies. *ReCALL, 25*(3), 321–339. <https://doi.org/DOI: 10.1017/S0958344013000141>

Описано один із складників трирічного дослідження використання технологій підсумкового оцінювання у старших класах середньої школи Австралії. Усне підсумкове оцінювання здійснювалося з використанням різних технологій: відео- й аудіозаписів, комп'ютеризованого тесту тощо. Результати останнього великою мірою збігаються з результатами звичних усних опитувань. Оцінювання усного мовлення онлайн є не менш ефективним, ніж за допомогою інших традиційних методів. Більше того, воно має певні переваги, як-то гнучкість доступу до тесту й можливість запису мовлення тестованого.

Crossley, S. & McNamara, D. (2013). Applications of text analysis tools for spoken response grading. *Language Learning & Technology, 17*(2), 171–192.

Досліджено потенціал автоматизованих індексів, пов'язаних з мовленням, використанням мови й розробленням тем для моделювання усного іншомовного мовлення під час тесту TOEFL. За допомогою покрокової лінійної регресії пояснено варіативність суджень щодо усномовленнєвих умінь. Автоматизовані індекси, пов'язані з вимірюванням типів слів, причинно-наслідковими зв'язками й лексичним репертуаром, дали змогу передбачити значний відсоток варіативності рейтингів зразків мовлення. Результати дослідження мають вагомe значення для розуміння природи усномовленнєвої компетентності й розроблення методів автоматизованого оцінювання.

Attali, Y., Lewis, W., & Steier, M. (2013). Scoring with the computer: Alternative procedures for improving the reliability of holistic essay scoring. *Language Testing, 30*(1), 125–141. <https://doi.org/10.1177/0265532212452396>

Експериментально оцінено кілька альтернативних процедур тестового оцінювання щодо підвищення їхньої надійності. Експериментальна перевірка базувалась на оцінюванні досвідченими експертами есе в межах тесту *GRE General Test*. Критерії передбачали міжособистісне погодження у середовищі експертів, узгодження з комп'ютеризованими оцінками й наскрізну надійність (*reliability*). Результати уможливили низку висновків: 1) використання модифікованих завдань, сконцентрованих на письмових вміннях вищого порядку, підвищило надійність одного типу завдань, але зменшило її для інших; 2) результати оцінювання груп аналогічних за обсягом есе не виявили впливу на загальне оцінювання; 3) оцінювання з використанням автоматизованих оцінок підвищує надійність експертних оцінок; 4) використання 18-бальної шкали оцінювання значно підвищило надійність тесту.

Shin, S. (2013). Developing a framework for using E-portfolios as a research and assessment tool. *ReCALL*, 25(3), 359–372. <https://doi.org/DOI: 10.1017/S0958344013000189>

Розглянуто ключові обмеження наявних електронних портфоліо, використовуваних як засіб оцінювання й досліджень. Стверджується, що попередні дослідження зазвичай спиралися на виявлення відмінностей між двома групами, а не внутрішньо особистісні відмінності у різних ситуаціях. Це не давало змоги дослідникам виявити індивідуальні тенденції розвитку учня, зокрема фактори ситуаційного впливу на його мовлення. Електронні портфоліо здатні подолати цю й інші проблеми. Запропоновано низку рекомендацій і пропозицій щодо усунення таких проблемних моментів й окреслено напрямки подальших досліджень.

#### 2014

Todd, R. (2014). Support Adaptive Testing: The effects of scaffolds in computer-based tests. *CALL-EJ*, 15(1), 1-20. Retrieved from [http://caliej.org/journal/15-1/Todd\\_2014.pdf](http://caliej.org/journal/15-1/Todd_2014.pdf)

Розглянуто можливості інноваційного підходу до комп'ютеризованого адаптивного тестування – *Support Adaptive Testing*, який базується на принципах розвивального навчання й навчальної автономії. Це звичайні комп'ютеризовані тести множинного вибору для перевірки прочитаного. Проте після виконання тесту студенти додатково працюють з помилковими відповідями, отримуючи зворотний зв'язок. Експериментально підтверджено позитивне ставлення студентів до адаптивного тестування.

Barkaoui, K. (2014). Examining the impact of L2 proficiency and keyboarding skills on scores on TOEFL-iBT writing tasks. *Language Testing*, 31(2), 241–259. <https://doi.org/10.1177/0265532213509810>

Досліджено вплив навичок набору на клавіатурі на результати тесту *TOEFL-iBT* (частина Writing). Результати показали, що загальний рівень володіння англійською мовою й умінь письма значною мірою сприяли відмінностям в отриманих результатах. При цьому навички набору на клавіатурі мали помітний, хоча й слабкий вплив на результати тесту. Крім того, цей вплив великою мірою залежав від типу завдань. Зроблено висновок про необхідність звернення уваги на проблему взаємозв'язку між навичками набору на клавіатурі, вміннями письма й процесом складання комп'ютеризованих мовних тестів.

#### 2015

Mozgalina, A., & Ryshina-Pankova, M. (2015). Meeting the challenges of curriculum construction and change: Revision and validity evaluation of a placement test. *Modern Language Journal*, 99(2), 346–370. <https://doi.org/10.1111/modl.12217>

Проаналізовано комп'ютеризовані тести в заповненні пропусків (*cloze*) як складника кваліфікаційних тестів (*placementtest*), використовувани в Джорджтаунському університеті (м. Вашингтон, США) в програмах вивчення німецької мови як іноземної. Наведено особливості розроблення й оцінювання таких тестів, що дають змогу підвищити якість навчальних програм і продемонструвати відмінності між тестованими з досить різними здібностями. Підкреслено вагомість контекстуально релевантних практик оцінювання для успіху програми, спрямованої на розвиток іншомовної грамотності.

Read, J. (2015). Issues in post-entry language assessment in English-medium universities. *Language Teaching*, 48(2), 217–234. <https://doi.org/DOI: 10.1017/S0261444813000190>

Розглянуто підґрунтя використання тестів *PELA* (*Post-Entry Language Assessments*) в Австралії. Зокрема, проаналізовано тест *DELNA* (*Diagnostic English Language Needs Assessment*), як частини процедур оцінювання іншомовних навичок й умінь в Оклендському університеті (Нова Зеландія). Частиною цих тестів є комп'ютеризоване тестування певних аспектів комунікативної компетентності. Проаналізовано такі ключові аспекти тестів як валідація, природа базової структури, “діагностичність”.

Chung, S. J., Haider, I., & Boyd, R. (2015). The English placement test at the University of Illinois at Urbana-Champaign. *Language Teaching*, 48(2), 284–287. <https://doi.org/DOI:10.1017/S0261444814000433>

Описано процедуру кваліфікаційних тестів (*placement test*), призначених для розподілу іноземних студентів на курси письма та/або вимови в Іллінойському університеті (Урбана-Шампейн). Одним з них є *Online English Placement Test*. Наведено історію тесту, основні статистичні дані й відмінності комп'ютеризованого тесту від його традиційного “паперового” аналогу.

Baturay, M. (2015) Online English language learners' perceptions of portfolio assessment. *Teaching English with Technology*, 15(4), 16-28. Retrieved from <http://www.tewtjournal.org/?wpdmact=process&did=ND11LmhvdGxpbms>

Дослідження проведено в межах онлайн-курсу англійської мови в одному з закладів вищої освіти Туреччини. Після впровадження портфоліо проаналізовано ставлення студентів до цього курсу як інструменту оцінювання. Студенти дуже позитивно ставляться до використання портфоліо, оскільки це допомагає їм побачити, як вони використовують виучувану мову. Студенти розмірковували щодо отриманих знань, відповідальності за навчання та мотивації до роботи над іноземною мовою в онлайн-режимі.

Poehner, M. E., Zhang, J., & Lu, X. (2015). Computerized dynamic assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing*, 32(3), 337–357. <https://doi.org/10.1177/0265532214560390>

Розглянуто проект розроблення онлайн-тестів множинного вибору для перевірки розуміння читання й аудіювання, що базуються на баченні медіації як невід'ємного складника мовної діагностики. Кожен елемент тесту супроводжується набором підказок – від прихованих до експліцитно виражених. У такий спосіб результати тестування демонструють не тільки правильність відповідей (фактичний результат), але й кількість підказок, отриманих під час тесту (опосередкований результат). Такий підхід дає змогу отримати ширшу картину розвитку комунікативної компетентності тестованих.

Levy, T., & Gertler, H. (2015). Harnessing technology to assess oral communication in business English. *Teaching English with Technology*, 15(4), 52–59.

Наведено підтверджувальні дані про кореляцію між використанням технології і залученістю студентів. Дані отримано шляхом створення й упровадження курсу, метою якого були мотивація, заохочення й оцінювання навичок усного мовлення. Студенти мали використовувати різні канали, зокрема відеозаписи, щоб продемонструвати свою комунікативну компетентність. Такий підхід дозволив студентам і викладачам долати проблему мовних класів великого наповнення.

Fageeh, A. I., & Fageeh; Abdulaziz I. (2015). EFL student and faculty perceptions of and attitudes towards online testing in the medium of Blackboard: Promises and challenges. *JALT CALL Journal*, 11(1), 41–62.

Досліджено ставлення студентів і викладачів до мовного оцінювання онлайн. Ставлення студентів опосередковано усвідомленням власної самоефективності, корисності тесту, задоволенням від виконання завдань тощо. Онлайн-тестування дає студентам змогу отримати безпосередній зворотний зв'язок й об'єктивні результати, що сприяє позитивному ставленню до такого способу оцінювання.

Tsurutani, C., & Imura, T. (2015). Designing online assignments for japanese language teaching. *JALT CALL Journal*, 11(3), 255–270.

Проаналізовано один із способів реалізації онлайн-оцінювання студентів-першокурсників японських університетів й обґрунтовано його використання. Розглянуто переваги та недоліки мовного оцінювання онлайн. Запропоновано шляхи ефективного реалізації такого оцінювання в інших навчальних курсах.

Cowie, N., & Sakui, K. (2015). Forum Assessment and e-learning: Current issues and future trends. *JALTCALL Journal*, 11(3), 271–281.

Стаття-огляд літератури, в якій зроблено спробу пояснити проблеми, що постають перед окремими викладачами й закладами освіти, які мають намір використати технології для контролю й оцінювання навчальних досягнень студентів. Розглянуто три основні тенденції комп'ютеризованого оцінювання: залучення засобів *Web 2.0*, розвиток спільних, командних проєктів і зростаюча роль цифрової грамотності.

Рубрику *Perspectives* номеру 99 (2) журналу *Modern Language Journal* присвячено вивченню іноземних мов онлайн. Деякі з поданих коментарів провідних фахівців галузі торкалися проблеми тестування в оцінювання в мережі Інтернет, наприклад:

Lin, C.-H. & Warschauer, M. (2015). Online foreign language education: What are the proficiency outcomes? *Modern Language Journal*, 99(2), 394–397. [https://doi.org/10.1111/modl.12234\\_1](https://doi.org/10.1111/modl.12234_1)

Поряд з іншими варіантами навчання іноземних мов онлайн досліджено онлайн-тести загального володіння мовою. Подано узагальнену характеристику тестів *Assessment of Performance toward Proficiency in Languages (AAPPL)*, *Versant*, *ACTFL Oral Proficiency Interview (OPI)*, *WebCape*, *Computer Assisted Screening Tool (CAST)* тощо.

Rubio, F. (2015). Assessment of oral proficiency in online language courses: Beyond reinventing the wheel. *Modern Language Journal*, 99(2), 405–408. [https://doi.org/10.1111/modl.12234\\_4](https://doi.org/10.1111/modl.12234_4)

Акцентовано увагу на проблемі поточного й підсумкового оцінювання онлайн, а також зворотного впливу тестування рівня загального володіння мовою на процес навчання. Так, можливості онлайн-середовища нівелюють бачення викладача як єдиного оцінювача. Взаємодія онлайн сприяє активнішій ролі учнів у процесі оцінювання й у такий спосіб розширює їхні можливості. При цьому способи оцінювання в межах онлайн-курсів мають відрізнитися від застосовуваних в аудиторних умовах.

Blake, R. (2015) The messy task of evaluating proficiency in online language courses. *Modern Language Journal*, 99(2), 408–412. [https://doi.org/10.1111/modl.12234\\_5](https://doi.org/10.1111/modl.12234_5)

Розглянуто різні способи оцінювання рівня загального володіння мовою в межах онлайн-курсів на основі стандартів ACTFL і Загальноєвропейських рекомендацій з мовної освіти. Вказано на обмеженість наявних засобів оцінювання, особливо для студентів початкового рівня володіння іноземною мовою, які стрімко набувають нових навичок і вмінь.

## 2016

Chapelle, C. A., & Voss, E. (2016). 20 years of technology and language assessment in Language Learning & Technology. *Language Learning & Technology*, 20(2), 116–128.

Огляд публікацій у журналі *Language Learning & Technology* (1997-2016 pp.) присвячено аналізу досліджень використання технологій під час оцінювання іншомовних навичок і вмінь.

Norris, J. M. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics*, 36, 230–244. <https://doi.org/10.1017/S0267190516000027>

Описано загальні процедури комунікативно орієнтованого оцінювання. Також зазначено, що використання комп'ютерних технологій для організації тестових завдань (віртуальна реальність тощо) разом з автоматизованими засобами оцінювання та надання зворотного зв'язку здатні створити тестове середовище, наближене до реального, яке уможливить якісніше оцінювання здібностей учня. Подальше використання технологій сприятиме лише збільшенню кількості автентичних комунікативних завдань у мовних тестах.

Chen, I. J. (2016). Hypertext glosses for foreign language reading comprehension and vocabulary acquisition: effects of assessment methods. *Computer Assisted Language Learning*, 29(2), 413–426. <https://doi.org/10.1080/09588221.2014.983935>

З'ясовано вплив чотирьох методів комп'ютеризованого оцінювання на засвоєння іншомовної лексики й розвиток умінь читання. З чотирьох обраних методів лише використання тесту



на основі спливаючих глосів не спричинило помітних змін рівня сформованості компетентностей. Найбільш ефективним виявився метод на основі використання завдань множинного вибору.

Baser, D., Koccha, T. J., & Ozden, M. Y. (2016). Developing a technological pedagogical content knowledge (TPACK) assessment for preservice teachers learning to teach English as a foreign language. *Computer Assisted Language Learning*, 29(4), 749–764. <https://doi.org/10.1080/09588221.2015.1047456>

Представлено процес розроблення й валідації самооцінювального опитування (TPACK-EFL), яке оцінює обсяг предметних технологічних і педагогічних знань майбутніх учителів англійської мови як іноземної. На початковому етапі за допомогою якісних методів дослідження (інтерв'ю з експертами й аналіз документів) розроблено елементи опитування, яке валідизовано шляхом двоетапного оцінювання факторної структури.

Burston, J., Neophytou, M. & Lamprianou, I. (2016). Developing and implementing a computer-adaptive test for English: The SLUPE experience. *CALL-EJ*, 17(1), 19-34. Retrieved from <http://callej.org/journal/17-1/Burston2016.pdf>

За допомогою безкоштовної інструментальної платформи *SLUPE* (Saint Louis University Placement Exam) розроблено комп'ютеризований адаптивний тест англійської мови, який був випробуваний як кваліфікаційний (*placement test*) протягом 2013-2014 рр. із залученням 1000 студентів Кіпрського університету технологій. Результати тесту порівняно з самооцінюванням студентів й оцінкою, наданою викладачами. Запропонований адаптивний тест найбільше корелював саме з викладацькою оцінкою, що дало змогу в подальшому запровадити його у практику кваліфікаційного тестування в університеті. Крім того, результати дослідження підтвердили можливість використання платформи *SLUPE* для розроблення адаптивних тестів.

Koyama, D., Sun, A., & Ockey, G. J. (2016). The effects of item preview on video-mediated MC listening assessments. *Language Learning & Technology*, 20(1), 148–165.

Розглянуто вплив форми й обсягу попереднього перегляду комп'ютеризованого тестового завдання на результати виконання тесту для перевірки аудитивних умінь. Обсяг попереднього перегляду завдань вплинув на результати тестування, але не вплинув на оцінки студентів з різним рівнем володіння мовою. Ідентифіковано елементи, які виявилися складнішими або простішими, ніж очікувалося. Вказано три можливих джерела такої ситуації: 1) частотність ввідних параметрів, 2) місце розташування ключової частини елемента, 3) наявність організаційних маркерів.

## 2017

Sasao, Y., & Webb, S. (2017). The Word Part Levels Test. *Language Teaching Research*, 21(1), 12–30. <https://doi.org/10.1177/1362168815586083>

Докладно досліджено комп'ютеризований тест “Word Part Levels Test” (WPLT) для перевірки знань щодо форми, значення й використання афіксів в англійській мові. Афікси дібрано на основі даних Британського національного корпусу. Також наведено особливості розроблення й упровадження комп'ютеризованого тесту.

Turner, R. L. (2017). The challenges of using the WebCAPE placement exam in an advanced Spanish grammar class. *Computer Assisted Language Learning*, 30(3–4), 247–258. <https://doi.org/10.1080/09588221.2017.1304424>

Здійснено спробу з'ясувати, чи придатний кваліфікаційний іспит WebCAPE, використовуваний для розподілу студентів університету на мовні курси відповідно їхньому рівню володіння мовою, для вимірювання навчальних досягнень на просунутому курсі навчання граматики. В іспиті використовуються технології комп'ютеризованого адаптивного мовного тестування, оптимізованого для студентів перших чотирьох семестрів навчання. Цей іспит успішно використано автором на просунутому курсі навчання іспанської граматики.

Tsai, S. C. (2017). Effectiveness of ESL students' performance by computational assessment and role of reading strategies in courseware-implemented business translation tasks. *Computer Assisted Language Learning*, 30(6), 474–487. <https://doi.org/10.1080/09588221.2017.1313744>

Досліджено ефективність перекладацької діяльності студентів на заняттях і використання ними стратегій читання у факультативному курсі англійської мови, призначеному для студентів старших курсів і виучуваному в обсязі 100 хвилин на тиждень протягом 12-ти тижнів. Автором проведено комунікативно орієнтоване навчання на базі комп'ютерної платформи. Результати експерименту продемонстрували ефективність запропонованої методики.

Wei, W., & Zheng, Y. (2017). An investigation of integrative and independent listening test tasks in a computerised academic English test. *Computer Assisted Language Learning*, 30(8), 864–883. <https://doi.org/10.1080/09588221.2017.1373131>

Здійснено всебічну оцінку й валідизацію аудитивної частини комп'ютеризованого тесту *Pearson Test of English Academic (PTE Academic)*, який складається з 11-ти типів тестових завдань, що дасть змогу оцінити вміння академічного аудіювання окремо або в поєднанні з іншими вміннями. Проаналізовано мету використання тестових завдань, а також оцінювальні вміння / конструкції й стимули завдань тесту *PTE Academic*. Сучасні технології дають змогу комп'ютеризованому тесту *PTE Academic* оцінити академічні здібності слухачів у режимі реального часу шляхом інтегрованого використання мультимодальних джерел.

Zhang, Z. (Victor). (2017). Student engagement with computer-generated feedback: a case study. *ELT Journal*, 71(3), 317–328. Retrieved from <http://dx.doi.org/10.1093/elt/ccw089>

Досліджено взаємодію студентів зі зворотним зв'язком, який надається системами автоматизованого оцінювання письма (*Automated Writing Evaluation, AWE*). Експериментально підтверджено позитивний вплив таких систем на формування іншомовної компетентності в письмі. Проте якісні показники впливу залежать від того, яким чином студенти взаємодіють зі зворотним зв'язком.

Tang, J., & Wu, Y. (2017). Impact of AWE rubrics and automated assessment on EFL writing Instruction. *International Journal of Computer-Assisted Language Learning and Teaching*, 7(2), 58–74. <https://doi.org/10.4018/IJCALLT.2017040104>

Досліджено вплив категорій системи автоматизованого оцінювання письма (*Automated Writing Evaluation, AWE*) *Writing Roadmap* у поєднанні з такою метакогнітивною діяльністю як автоматизоване оцінювання на навчання іншомовного письма. Виявлено, що здійснення поточного оцінювання з використанням категорій системи автоматизованого оцінювання письма сприяє досягненню балансу між цілями навчання й оцінювання, покращує взаємодію викладачів і студентів, розвиває навчальну автономію.

Jin, Y., & Wu, E. (2017). An argument-based approach to test fairness. *International Journal of Computer-Assisted Language Learning and Teaching*, 7(3), 58–72. <https://doi.org/10.4018/IJCALLT.2017070104>

Завдяки використанню інноваційних методик тестування розглянуто шляхи запобігання випадкам нечесних практик у мовному оцінюванні. Подано загальний опис нечесних практик під час тесту *College English Test (CET)*. Виявлено альтернативні форми забезпечення надійності під час використання різних версій і форм тесту. Розглянуто ефективність використання запропонованих підходів до емпіричної оцінки чесності тестів.

Вип. 14 (2) журналу *Language Assessment Quarterly* присвячено підсумковим кваліфікаційним екзаменам у Китаї. Деякі зі статей торкалися теми комп'ютерного тестування:

Jin, Y., & Yan, M. (2017). Computer literacy and the construct validity of a high-stakes computer-based writing assessment. *Language Assessment Quarterly*, 14(2), 101–119. <https://doi.org/10.1080/15434303.2016.1261293>

Досліджено, чи призвів перехід від “паперового” до комп'ютерного тестування в китайському підсумковому кваліфікаційному іспиті *College English Test* до варіативності, невідповідної

матеріалу, що оцінюється в тесті. Аналіз результатів, складності тексту й мовних помилок довів, що когнітивні процеси тестованих під час комп'ютеризованого письма були схожими на такі ж процеси під час традиційного тестування. Стверджується, що у 21-му столітті комп'ютерне мовне тестування має бути переосмислене: комп'ютерна грамотність повинна розглядатися як важливий контекстуальний аспект, який взаємодіє з матеріалом, вимірюваним комп'ютерним тестом.

He, L., & Min, S. (2017). Development and validation of a computer adaptive EFL test. *Language Assessment Quarterly*, 14(2), 160–176. <https://doi.org/10.1080/15434303.2016.1162793>

Описано комп'ютерний адаптивний тест з англійської мови, що оцінює вміння англомовних аудіювання й читання. Докладно описано методи розроблення такого тесту, включаючи створення банку тестових завдань, вибір відповідних моделей теорії для калібрування тестових завдань, процедури вибору завдань тощо. Проаналізовано валідність розробленого тесту й запропоновано шляхи його використання.

Yu, G., & Zhang, J. (2017). Computer-based English language testing in China: Present and future. *Language Assessment Quarterly*, 14(2), 177–188. <https://doi.org/10.1080/15434303.2017.1303704>

В авторському огляді досліджень комп'ютерного тестування англійської мови, здійснених китайськими вченими й опублікованих у китайських академічних журналах в останні десятиліття, виявлено основні наукові напрямки, проблеми й тенденції, розуміння яких сприятиме кращому використанню комп'ютерних технологій у мовному тестуванні.

#### 2018

McNeil, L. (2018). Understanding and addressing the challenges of learning computer-mediated Dynamic Assessment: A teacher education study. *Language Teaching Research*, 22(3), 289–309. <https://doi.org/10.1177/1362168816668675>

Дослідження спрямоване на виявлення й розуміння складнощів, які можуть виникати під час опрацювання майбутніми та працюючими вчителями навчального курсу, що містить комп'ютеризоване динамічне оцінювання. Результати першого циклу експерименту продемонстрували, що вчителі мали проблеми з використанням стратегій, пов'язаних із системою підтримки в межах курсу та розумінням значення поняття динамічного оцінювання.

Ebadi, S., Weisi, H., Monkaresi, H., & Bahramlou, K. (2018). Exploring lexical inferencing as a vocabulary acquisition strategy through computerized dynamic assessment and static assessment. *Computer Assisted Language Learning*, 31(7), 790–817. <https://doi.org/10.1080/09588221.2018.1451344>

Проаналізовано результати квазіекспериментального дослідження, в якому порівняно ефективність комп'ютеризованого динамічного оцінювання на противагу статичному оцінюванню щодо розпізнавання нових лексичних одиниць. Отримані результати свідчать, що правильно організоване розпізнавання незнайомих слів разом з процедурами динамічного оцінювання здатні збільшити приріст словникового запасу студентів.

Park, M. (2018). Innovative assessment of aviation English in a virtual world: Windows into cognitive and metacognitive strategies. *ReCALL*, 30(2), 196–213. <https://doi.org/DOI: 10.1017/S0958344017000362>

Досліджено можливості використання вербальних даних, отриманих від тестованих під час комп'ютеризованого тестування. Шукаються відповіді на запитання: 1) Які стратегії використовуються під час виконання віртуального завдання? та 2) Як можна інтерпретувати стратегії у порівнянні з діяльністю тестованого? Виявлено позитивний зв'язок між кількістю задіяних когнітивних / метакогнітивних стратегій і результатами тесту. Подані у віртуальному середовищі специфічні завдання здатні надавати учню можливість працювати з автентичними матеріалами та розвивати стратегічну й мовні компетентності.

Long, A. Y., Shin, S.-Y., Geeslin, K., & Willis, E. W. (2018). Does the test work? Evaluating a web-based language placement test. *Language Learning & Technology*, 22(1), 137–156.

У контексті наявних підходів до оцінювання розглянуто можливі шляхи інтерпретації й використання результатів університетського кваліфікаційного тесту (*placement test*) з іспанської мови. Розроблено комп'ютеризований кваліфікаційний тест. Проаналізовано його внутрішню структуру й валідність щодо функціональності, змісту й надійності тесту. Надано рекомендації для викладачів і розробників тестів.

Kamrood, A., Davoudi, M., Amirian, S. & Ghaniabadi, S. (2018). Transcendence of learning in an online computerized dynamic test of English listening. *CALL-EJ*, 19(1), 23-42. Retrieved from <http://callej.org/journal/19-1/Kamrood-Davoudi-Amirian-Ghaniabadi2018.pdf>

Досліджено трансцендентність навчання в умовах розроблення онлайн-комп'ютеризованого динамічного тестування англомовного аудіювання. За допомогою графічного відображення та кореляції порівняно результати динамічного й трансцендентного оцінювання. Доведено, що викладачі мови можуть точніше й інформативніше передбачати розвиток учнів завдяки аналізу й порівнянню результатів різних типів оцінювання.

Lesnov, R. O. (2018). Content-rich versus content-deficient video-based visuals in L2 academic listening tests. *International Journal of Computer-Assisted Language Learning and Teaching*, 8(1), 15–30. <https://doi.org/10.4018/ijcallt.2018010102>

Порівняно виконання завдань тестування вмінь академічного аудіювання у двох режимах: з використанням лише аудіо- та з поєднанням аудіо- й відеоматеріалів. Розроблено й перевірено новий метод класифікації відео-матеріалів, в основі якого лежить розподіл у континуумі від змістовно обмежених (сприйняття яких не дозволяє повноцінно виконати завдання на перевірку розуміння) до змістовно наповнених (сприйняття яких дає змогу повноцінно виконати завдання на перевірку розуміння).

Quaid, E. D. (2018). Output register parallelism in an identical direct and semi-direct speaking test. *International Journal of Computer-Assisted Language Learning and Teaching*, 8(2), 75–91. <https://doi.org/10.4018/IJCALLT.2018040105>

Порівняно результати комп'ютеризованого тесту для перевірки вмінь усного мовлення *AptisGeneral* й безпосереднього усного опитування. В межах тематичного аналізу досліджено вихідні характеристики, виявлені в попередніх дослідженнях. З'ясовано, що результати, отримані під час комп'ютеризованого тесту, були менш контекстуалізовані, з мінімально вищою лексичною щільністю і синтаксичною складністю. Запропоновано пояснення таких результатів і можливі шляхи їх використання в практиці оцінювання усного мовлення.

Lee, S., & Winke, P. (2018). Young learners' response processes when taking computerized tasks for speaking assessment. *Language Testing*, 35(2), 239–269. <https://doi.org/10.1177/0265532217704009>

Досліджено різні аспекти діяльності молодших школярів (8-10 років) під час виконання комп'ютеризованого тестування усномовленнєвих вмінь. Досліджено концентрацію уваги дітей щодо різних складників тестових завдань (наприклад, підказках, малюнках, таймері) за допомогою фіксації зору руху очей. Експериментально перевірено цілком очікувану помітну різницю в результатах носіїв і неносіїв мови. Крім того, обидві групи продемонстрували відмінності в концентрації уваги під час тестування.

van Hout, R., Lobo, V., McQueen, JM., Unsworth, S., Goriot, C., & Broersma, M. (2018). Using the peabody picture vocabulary test in L2 children and adolescents: effects of L1. *International Journal of Bilingual Education and Bilingualism*, 1-23. <https://doi.org/10.1080/13670050.2018.1494131>

Досліджено, чи можна вважати комп'ютеризований тест *Peabody Picture Vocabulary Test (PPVT-4)* надійним інструментом для перевірки англомовної лексичної компетентності, а також в який спосіб знання рідної мови впливають на результати тесту. Експериментально перевірено в умовах середньої школи, що тест *PPVT-4* не є надійним для всіх категорій учнів. Крім того, тестовані демонстрували результати в тих завданнях, де спостерігалася фонетична спорідненість рідної й іноземної мов.

2019

Hockly, N. (2019). Automated writing evaluation. *ELT Journal*, 73(1), 82–88. <https://doi.org/10.1093/elt/ccy044>

Розглянуто сучасний стан розвитку систем автоматизованого оцінювання письма (*Automated Writing Evaluation, AWE*). Проаналізовано його переваги й недоліки. Акцентовано увагу на недостатній ефективності систем автоматизованого оцінювання письма та їх обмеженості в контексті потреб учителів й учнів. Зазначено, що за умови ретельного розроблення системи автоматизованого оцінювання письма можуть бути достатньо корисними для навчання й наукових досліджень.

Khoshshima, H., Morteza, S., Toroujeni, H., Thompson, N., & Ebrahimi, M. R. (2019). Computer-based (CBT) vs. paper-based (PBT) testing: Mode effect, relationship between computer familiarity, attitudes, aversion and mode preference with CBT test scores in an Asian private EFL context. *Teaching English with Technology*, 19(1), 86–101. Retrieved from <http://www.tewtjournal.org/?wpdmact=process&did=NTYxLmhvdGxpbnms>

Здійснено спробу виявити наявність розбіжностей між результатами, які іранські учні отримували під час комп'ютерного й “паперового” тестувань. Експериментальне навчання не продемонструвало значної відмінності результатів учнів у двох режимах. Крім того, не виявлено статистично значущої кореляції між ставленням учнів (як позитивним, так і негативним) і результатами тестування.

Solano-Flores, G., Chia, M., & Kachchaf, R. (2019). Design and use of pop-up illustration glossaries as accessibility resources for second language learners in computer-administered tests in a large-scale assessment system. *International Multilingual Research Journal*, 1–17. <https://doi.org/10.1080/19313152.2019.1611338>

Прозвітовано про довготривалий проект зі створення візуально доступних ресурсів як частини великої комп'ютеризованої системи оцінювання. Детально описано процес розробленняпливаючих візуальних глосаріїв, призначених для пояснення незнайомих лексичних одиниць. Розглянуто, в який спосіб такі візуальні елементи сприяють індивідуалізації навчання, а також валідному й чесному оцінюванню.

Mizumoto, A., Sasao, Y., & Webb, S. A. (2019). Developing and evaluating a computerized adaptive testing version of the Word Part Levels Test. *Language Testing*, 36(1), 101–123. <https://doi.org/10.1177/0265532217725776>

Прозвітовано про розроблення й оцінювання засобу комп'ютеризованого адаптивного тестування в межах тесту *Word Part Levels Test (WPLT)*. Засіб був розроблений для максимального підвищення ефективності *WPLT* як діагностичного тесту. Результати продемонстрували, що засіб може надати тестованим діагностичну інформацію про знання афіксів англійської мови з такою ж або більшою точністю, ніж попередні версії *WPLT*. Обговорено педагогічні умови використання засобу й загальні проблеми комп'ютеризованого адаптивного тестування.