

Національний тест з іноземної мови: ознаки і перспективи розробки

У контексті приєднання України до Болонського процесу та Європейського мовного простору реформування системи освіти у нашій країні є конче необхідним. Це стосується не лише вищої школи, яка переходить до модульно-рейтингової організації навчання та оцінювання. Корінні зміни започатковані й у загальноосвітніх навчальних закладах в організації випускних іспитів, які проводитимуться у формі зовнішнього незалежного тестування і, як у багатьох європейських країнах та США, будуть вважатися як вступні до вищих закладів освіти.

Влітку 2006 року зовнішнє тестування вже було впроваджене, але лише в деяких школах та за бажанням учнів з таких предметів як українська мова, історія України і математика. Наступного року зовнішнім тестуванням будуть охоплені всі випускники країни. Розшириться також коло предметів, знання з яких будуть перевірятися за допомогою тестів. Виняток буде зроблено стосовно предмета „Іноземна мова” (ІМ), що обумовлене як його специфікою, так і закономірностями розробки тестів з ІМ. Яким бути зовнішньому тестуванню з ІМ, які проблеми мають вирішити освітяни і суспільство для його успішного впровадження – висвітлення цих та деяких інших питань є метою запропонованої статті.

Сьогодні, коли вже не потрібно нікого переконувати в перевагах володіння однією чи двома ІМ, суспільство зацікавлене в тому, щоб мовна освіта в Україні поряд з високою ефективністю набула також якості на рівні міжнародних стандартів, що надало б можливості громадянам країни відчувати себе повноправними членами європейського співтовариства. Для досягнення цієї мети за основу розробки Програм для середньої, вищої мовної та немовної шкіл Україною прийняті „Загальноєвропейські рекомендації з мовної освіти: вивчення, викладання, оцінювання” (ЗЕР)¹. Аналізуючи

ЗЕР, Н.Ф.Бориско, зокрема, зазначає, що вони мають усі ознаки енциклопедії (хоча й не є такою) в галузі навчання ІМ, а також наголошує на нормативному (а не рекомендаційному) характері документа, внаслідок чого він може служити досить надійною базою для гармонізації мовної освіти в різних країнах, а саме – стандартизації вимог до її якості, прозорості та співвіднесеності результатів навчання ІМ як в національному, так і в міжнародному вимірах².

Отже, згідно з Державним стандартом і Програмою для загальноосвітніх шкіл випускники мають оволодіти іноземною мовою на рівні B1+.³ Проте не викликає сумніву той факт, що далеко не всі випускники України з різних причин (природні здібності – не остання з них) будуть спроможні продемонструвати такий високий рівень володіння іншомовними вміннями. Водночас у країні існує багато шкіл з поглибленим вивченням ІМ, гімназій та ліцеїв, які забезпечують вищий рівень володіння ІМ, ніж загальноприйнятий B1+. Це означає, що й тест для остаточного оцінювання рівня навченості з ІМ випускників усіх шкіл України повинен бути багаторівневим, що значно ускладнює завдання тестологів, які розробляють зміст випускного іспиту з ІМ.

Вітчизняними методистами накопичений досить великий досвід використання тестових технологій для оцінювання досягнень тих, хто навчається⁴, проте питання розробки незалежного тесту в загальнонаціональному масштабі ще не порушувалося. Тому з лютого 2006 року за участю Міністерства освіти і науки України і Британської Ради в Україні розпочалася робота Всеукраїнського проекту – підготовка як самих тестів, так і спеціалістів, які мають науково обґрунтувати, розробити і впровадити ці тести.

Отже, що необхідно зробити для того, щоб запропонувати суспільству якісний та надійний тест з ІМ?

¹ Загальноєвропейські Рекомендації з мовної освіти : вивчення, викладання, оцінювання / Наук. редактор укр. видання доктор пед. наук, проф. С.Ю. Ніколаєва. – К.: Ленвіт, 2003. – 273 с.; Програма з англійської мови для університетів/інститутів (п'ятирічний курс навчання): Проект / Кол. авторів: С.Ю. Ніколаєва, М.І. Соловей (керівники), Ю.В. Головач та ін. – Вінниця: Нова книга, 2001. – 245 с.; Бориско Н.Ф., Гутник В.М. и др. Типовая программа по практике устной и письменной речи для педагогических факультетов университетов и педагогических институтов (немецкий язык) / под ред. Н.Ф. Бориско. – Киев: Ленвіт, 2004. – 254 с.; Програма з англійської мови для професійного спілкування. Кол. авторів: Г.Є. Бакаєва, О.А. Борисенко та ін. – К.: Ленвіт, 2005. – 119 с.; Державний стандарт базової і повної середньої освіти (іноземні мови) // Іноземні мови. – 2004. – №1. – С. 3-7.

² Бориско Н.Ф. «Общеввропейские компетенции владения иностранным языком: Изучение, обучение, оценка». Анализ некоторых аспектов // Іноземні мови. – 2005. – № 1. – С. 8-14.

³ Конкретизований зміст умінь учнів у продуктивних, рецептивних та інтерактивних видах мовленнєвої діяльності, а також стратегій і компетенцій поданий у „Загальноєвропейських рекомендаціях з мовної освіти...” – Ленвіт, 2003 і в численних публікаціях (див., наприклад, Шаленко О. Це раз про рівні володіння іноземною мовою // Іноземні мови. – 2004. – №2. – С.3-8; Яценко Л.М. Вчителям про основні положення Загальноєвропейських рекомендацій з мовної освіти // Іноземні мови. – 2004. – №2. – С.57-59).

⁴ Ніколаєва С.Ю., Бадаєв Г.С. Тестові завдання з англійської мови / Tasks for Language Testing. – К.: Освіта, 1998. – 216 с.; Петрашук О.П. Вплив вікових психологічних особливостей учнів на їх готовність до виконання тесту з іноземної мови // Іноземні мови. – 2000. – № 2. – С. 3-4.; Петрашук О.П. Тестові завдання з читання для рубіжного / підсумкового контролю у навчанні англійської мови в середній загальноосвітній школі // Іноземні мови. – 2000. – №1. – С. 20-23.; Петрашук О.П. Тестовий контроль у навчанні письма учнів середньої загальноосвітньої школи // Іноземні мови. – 1999. – № 3. – С. 11-14.; Петрашук О.П. Типи тестових завдань для контролю іншомовного говоріння учнів середньої загальноосвітньої школи // Іноземні мови. – 1999. – №1. – С. 14-18.

ЗАПРОШУЄМО ДО ДИСКУСІЇ

Передусім пригадаємо теорію і дамо визначення основним поняттям тестології¹ як науки.

Тестування (*testing*) як термін зустрічається в літературі у вузькому смислі – у значенні використання і проведення тесту – і в широкому смислі – як сукупність процедурних етапів планування, складання та випробування тестів, обробки й інтерпретації їх результатів. У тестуванні в широкому смислі, як правило, виділяють три основні стадії:

I стадія – розробка передтесту: планування тесту, складання передтесту, апробація передтесту, інтерпретація результатів передтесту;

II стадія – розробка остаточного тесту: перепланування передтесту, складання та оформлення остаточного варіанта тесту, підготовка специфікацій та інструкцій;

III стадія – реалізація тесту: підготовка до застосування тесту, проведення тесту, обробка результатів тесту, інтерпретація результатів тесту.

Розробка та застосування мовних і мовленнєвих тестів є сферою компетенції лінгводидактичного тестування (*language testing*). "...Під лінгво-дидактичним тестом розуміють підготовлений згідно з певними вимогами комплекс завдань, які пройшли попереднє випробування з метою визначення показників його якості і дозволяють виявити в учасників тесту² ступінь їхньої мовної (лінгвістичної) та/чи мовленнєвої (комунікативної) компетенції, і результати якого піддаються певному оцінюванню за задалегідь встановленими критеріями"³.

Тести з великою кількістю завдань (більше 40-50) називаються комплексними тестами або тестовими батареями. Вони складаються з частин (*parts*) і субтестів (*subtests*).

За метою застосування тести розподіляються на констатуючі, діагностичні та прогностичні. Констатуючі тести включають тести навчальних досягнень і тести загального володіння іноземною мовою.

Тести навчальних досягнень (*achievement tests*) складаються у точній відповідності до програми та підручника або засвоєного навчального матеріалу і використовуються для здійснення поточного, рубіжного і підсумкового контролю.

Тести загального володіння іноземною мовою (*proficiency tests*) складаються незалежно від програм і підручників попереднього навчання. У тест такого типу може бути включена перевірка тих іноземних навичок і вмінь, які будуть необхідні у подальшому навчанні або роботі. В залежності від конкретної мети їх використання вони можуть застосовуватися для відбору абітурієнтів (*selection tests*), розподілу учнів / студентів по групах (*place-*

ment tests), уточнення програми навчання окремої групи учнів / студентів (*planning instruction tests*).

За статусом контролюючої програми виділяють стандартизовані і нестандартизовані тести.

Стандартизованим тестом (*standardized test*) називають такий тест, який пройшов апробацію на досить великій (500-1000 чоловік) кількості учасників тестування і який має стабільні й допустимі показники якості, а також специфікацію – паспорт з нормами, умовами та інструкціями для його багаторазового використання в різних умовах.

Нестандартизованими називаються викладацькі/вчительські (*classroom*) тести, які складаються самими викладачами / вчителями для своїх студентів/учнів. Їх називають по-різному: тест, тестова контрольна робота, квіз (*quiz*). Ці тести короткі і для них не є обов'язковим визначення усіх показників якості, однак вони проходять хоча б одну апробацію з метою визначення складності завдань.

Проміжним є локально стандартизований тест, показовою вибіркою в якому є контингент тестованих однієї чи кількох шкіл, одного чи кількох вищих навчальних закладів міста (100-300 осіб).

Отже, зважаючи на вищевикладене, для зовнішнього тестування випускників українських шкіл з ІМ ми маємо розробити багаторівневий стандартизований тест загального володіння іноземною мовою (*proficiency test*). Оскільки від точності й об'єктивності вимірювання досягнень учнів будуть значною мірою залежати навчальні та професійні перспективи кожного окремого випускника школи, при розробці цього відповідального тесту (*high-stakes test*) ми повинні ретельно проаналізувати і врахувати параметри його якості.

Якість тесту визначається насамперед його *придатністю* (*test usefulness*), яка розглядається нами, слідом за Л.Бахманом та А.Палмером, як базова концептуальна послідовність тестології. Загальна формула придатності мовного тесту виглядає так:

Функціональність (*Usefulness*) =
надійність (*reliability*) + валідність конструкту (*construct validity*) + автентичність (*authenticity*) +
інтерактивність (*interactiveness*) + вплив (*impact*) +
практичність (*practicality*)⁴.

Розглянемо складники придатності мовного тесту докладніше.

Перш за все, тест з ІМ має бути *надійним інструментом* вимірювання певного вміння (умінь) незалежно від специфіки його застосування у контексті різноманітних тестових ситуацій. Наприклад, у разі розробки декіль-

¹ Ніколаєва С. Ю., Петрашук О. П. Тестовий контроль лексичних навичок аудіювання (англійська мова). – К.: Ленвіт, 1997 – С. 4-5.

² У подальшому замість загальноприйнятого у вітчизняній фаховій літературі терміна „учасники тесту” ми будемо вживати термін „кандидат” як більш адекватний, на нашу думку, корелят відповідних англійських слів *test-taker*, *testee* і *candidate*. Вважаємо, що лексема „учасник тесту” вміщує більш широкий обсяг поняття, оскільки його денотативна основа охоплює не лише осіб, що складають певний тест (*test-takers* і *testees*), але й тих, хто бере участь у „житті” тесту на різних етапах тестового циклу та адміністраторів, тих, хто тест проводить (*testers*); усіх тих, кого залучено у тестовий процес прямо або опосередковано – вчителів, батьків, державні установи (*stakeholders*) тощо.

³ Коккота В. А. Лингво-дидактическое тестирование. – М.: Высшая школа, 1989. – С. 8.

⁴ Bachman, L.F. and A. Palmer. *Language Testing in Practice*. – Oxford: Oxford University Press, 2000. – p.18.

ЗАПРОШУЄМО ДО ДИСКУСІЇ

кох варіантів одного й того ж мовного тесту для того, хто його складає, немає ніякої різниці, який із запропонованих варіантів виконувати: кандидат повинен отримати абсолютно ідентичну кількість балів незалежно від варіанта, оскільки усі тестові завдання в них були розроблені за єдиними вимогами, з урахуванням єдиних цілей і передбачають єдині критерії оцінювання, хоча й побудовані на різному мовному матеріалі.

Надійність як *необхідна* ознака будь-якого тесту з ІМ не є, однак, *достатньою* передумовою придатності останнього. Уявімо, наприклад, ситуацію, коли нам потрібно провести тест з метою розподілу учнів/студентів на підгрупи при проведенні курсу ІМ (*placement test*). Граматичний тест, розроблений, скажімо, у форматі "multiple choice", може надати дійсно достовірні та надійні результати. Але наскільки обґрунтованими слід вважати ці результати для оцінки загального рівня сформованості мовних (у широкому розумінні) навичок кандидатів, беручи до уваги той факт, що знання граматики – це лише один з багатьох параметрів, які потребують вимірювання у контексті зазначеного типу тестів? Цілком очевидно, що оцінюванню підлягають і такі аспекти як рівень сформованості мовленнєвої, соціокультурної та прагматичної компетенції кандидатів; ступінь зрілості у залученні ними певних когнітивних стратегій до виконання тестових завдань тощо.

У цьому випадку йдеться про *валідність конструкту*, який покладено в основу концепції того чи іншого лінгводидактичного тесту. Сам конструкт (*construct*) тлумачиться в тестології як психологічний концепт, специфічне визначення вміння, яке покладене в основу певного тесту та інтерпретації його результатів¹. Відповідно, кожен тест покликаний здійснити вимір одного чи декількох конструктів².

У контексті зазначених теоретичних посилань *валідність конструкту* – однієї з базових функціональних ознак лінгводидактичного тесту – слід визначити як обґрунтованість (*meaningfulness*) та адекватність (*appropriateness*) висновків, які ми робимо, спираючись на результати тестування. Інакше кажучи, *інтерпретуючи* останні як показники оцінюваних мовних здібностей (*language ability*) кандидатів, ми повинні завжди чітко визначитися з відповіддю на ключове запитання: „Якою мірою ми можемо обґрунтувати наші висновки?“. Єдиний спосіб довести „валідність“ тесту – це надати доказ того, що показники виконання тесту відображають вимір саме тих умінь, які ми намагались оцінити. Але для того щоб надати такі докази, ми маємо визначити власне „конструкт“ – теоретичне підґрунтя вимірюваних параметрів.

Для того щоб довести придатність і доцільність використання певного лінгводидактичного тесту, слід також звернути увагу ще на один важливий його показ-

ник, а саме – ступінь відповідності завдань, з яких він складається, завданням реального використання іноземної мови у ситуаціях спілкування. Саме цей показник, який визначається у фаховій літературі як *автентичність* (*authenticity*) тесту з ІМ, не завжди чітко прослідковується в наявних навчальних посібниках та збірках завдань для тестування. Проте саме він повинен стати основою розробки і впровадження у практику оцінювання в Україні лінгводидактичних тестів якісно нового, європейського зразка.

Ще одна базова ознака тесту, якою ми досить часто нехтуємо, – це його *інтерактивність* (*interactiveness*), тобто ступінь відповідності кандидата до виконання тестового завдання. Адже не викликає сумніву той факт, що саме особистісні характеристики значною мірою обумовлюють потреби і способи використання індивідом мови (зокрема іноземної) як засобу спілкування та вираження власних думок. Відповідно, згадані параметри кандидатів повинні також враховуватися під час розробки тестів з ІМ та проведенні самого тестування. Звичайно, несподівану зміну настрою учня або його втому (фізичну чи моральну) важко передбачити повною мірою. Однак, вважають Л.Бахман і А.Палмер, існують чотири блоки індивідуальних особливостей кандидатів, вплив яких на показники виконання ними тесту слід брати до уваги. Це, насамперед, такі особистісні ознаки як вік, стать, національність, освітній рівень, рідна мова, ступінь підготовленості та попередній досвід кандидата щодо виконання запропонованого йому тесту³. Не менш важливими у контексті, що розглядається, виявляються також фонові знання (*topical knowledge*) та емоційна сфера (*affective schemata*) того, хто тестується; так само як його мовні здібності (*language ability*), включаючи мовну компетенцію/мовні знання (*language competence / language knowledge*) і стратегічну компетенцію (*strategic competence*) – набір свідомо виконуваних і контрольованих операцій (*metacognitive strategies*).

Наступна риса, що свідчить на користь придатності мовного тесту, це *вплив* (*impact*), який останній має на суспільство та освітню систему в цілому (макро-вплив / *macro-impact*) і на окремих індивідів у межах цієї системи – самих кандидатів, учителів, батьків тощо (мікро-вплив / *micro-impact*). Багато аспектів саме цієї ознаки тесту залишаються нез'ясованими у вітчизняному контексті. Брак досвіду широкомасштабного тестування, зокрема тестування на загальнонаціональному рівні; нерозвиненість системи „зворотного зв'язку“ (*feedback*) і практики публічного звітування про результати проведеного тестування; недостатня інформованість усіх учасників процесу тестування про їх права та обов'язки; питання про те, яким чином результати тестування повинні впливати на процес практичного викладання ІМ (так званий „*washback*“) – ось далеко не

¹ „A construct is a psychological concept, which derives from a theory of the ability to be tested“ – [Alderson, Ch.J. *Assessing Reading*. – Cambridge: Cambridge University Press, 2005. – p.118]; „... we consider a construct to be the specific definition of an ability that provides the basis for a given test or test task and for interpreting scores derived from this task“ [Bachman, L.F. and A. Palmer. *Language Testing in Practice*. – Oxford: Oxford University Press, 2000. – p.21].

² Alderson, Ch.J. *Assessing Reading*. Cambridge: Cambridge University Press, 2005. – p.118.

³ Bachman, L.F. and A. Palmer. *Language Testing in Practice*. Oxford: Oxford University Press, 2000. – p. 64-66.

ЗАПРОШУЄМО ДО ДИСКУСІЇ

вичерпаний список тих проблем, з якими найближчим часом доведеться зустрітися як фахівцям-тестологам, так і учням / студентам, а також службовцям у владних кабінетах різних рівнів.

Коментуючи останню із базових характеристик, які надають лігводидактичному тесту ознак функціональності, слід зазначити, насамперед, що за своєю природою вона суттєво відрізняється від згаданих вище ознак. Так, усі розглянуті особливості стосуються інтерпретації та використання результатів тестування, а *практичність (practicality)* тесту регулює умови, в яких цей тест може бути впроваджений. Отже основним концептуальним посланням у цьому сенсі є думка про те, що ресурси (людські й матеріальні), залучені до процесу розробки, вдосконалення та використання тесту, не повинні перевищувати можливостей, на які спільнота розраховує. В цілому практичність – це справа того, якою мірою вимоги, зазначені у специфікації до відповідного тесту, можна задовольнити з урахуванням наявних у певному соціальному контексті лімітів ресурсів.

Отже, як бачимо, поняття „придатність тесту” не тільки синтезує в собі найважливіші ознаки якісного тестового продукту (валідність, надійність, автентичність). В ньому опосередковано окреслені стадії тестування (*testing cycle*), на яких ми зупинялися вище, і наголошено на активній та відповідальній участі спільноти у спостереженні за перебігом процесу розробки і впровадження тесту. Тому першим кроком для забезпечення співпраці розробників тесту, адміністративних органів та суспільства передумовою прозорості майбутнього іспиту повинна бути розробка специфікації тесту. Остання вже згадувалася нами як паспорт тесту, в якому визначені норми, умови та інструкції для його багаторазового використання. Далі розглянемо сутність специфікації докладніше.

Специфікація – це офіційний документ, в якому визначено, що саме має виміряти тест і в який спосіб він це вимірюватиме¹. Специфікація є своєрідним „планом роботи” для тих, хто розробляє тест (*test designers*), оскільки в ній чітко визначено, для кого передбачений цей тест; з якою метою здійснюється вимір знань, умінь тощо; які саме знання, вміння та в якому обсязі мають бути перевірені; скільки частин (*papers, sections*) тест має містити для перевірки того, що планується перевірити; які способи перевірки доцільно застосовувати; скільки часу потрібно для виконання кожної частини і загалом усього тесту тощо. Специфікація необхідна також для тих, хто редагує та вдосконалює те, що розроблено укладачами тесту (*test editors and moderators*). Специфікація такого характеру є суворо конфіденційним документом, призначеним суто для службового використання.

Дещо в іншому, скороченому обсязі, але специфікація потрібна й тим, хто зацікавлений у встановленні валідності тесту. Це можуть бути незалежні інститути або особи, яких більше цікавить не детальна інформація про

кількість частин або тривалість тесту, а теоретичне обґрунтування формату тесту, тобто чому він має саме такий, а не інший вигляд.

Учасники тестування – власне кандидати і вчителі, що їх навчають, – також повинні мати свою специфікацію тесту. Для них важливо буде знати зміст тесту, тобто які іншомовні вміння будуть перевірятися за допомогою тесту, які типи тестових завдань (*test tasks and test items*) передбачені для встановлення рівня володіння різними видами мовленнєвої діяльності, які бали отримуватимуть кандидати за правильні відповіді і яку кількість балів необхідно отримати для успішного складання тесту.

Для забезпечення якісної підготовки до тесту окремою версією специфікації мають бути забезпечені й автори та видавці посібників, збірок тестів і довідкових видань. Вони повинні насамперед усвідомлювати, для якого рівня володіння ІМ видаються ті чи інші книжки, наскільки повно їх зміст відповідає вимогам до вмінь, що підлягатимуть перевірці на іспиті.

Важливість розробки досконалої специфікації незалежного тесту з ІМ для українських випускників, на наш погляд, не викликає сумніву. Її появі передують ретельне вивчення і критичне осмислення як теоретичних джерел з тестології, так і результатів розробки та впровадження аналогічних іспитів у декількох країнах Європи. Досвід європейських колег свідчить: для того щоб служити надійним інструментом вимірювання знань та вмінь, тест має бути розроблений на високому професійному рівні, а для цього тестологи повинні дотримуватися усіх стадій підготовки тесту – від дослідження реальної картини успішності навчання ІМ в сучасній школі, пілотування завдань майбутнього тесту до підготовки великої кількості незалежних екзаменаторів (*markers and scorers*), які забезпечать об'єктивне оцінювання.

У цілому, підсумовуючи наведені вище думки, слід зазначити таке: знаходячись на самому початку довгого й відповідального шляху до впровадження зовнішнього тестування в Україні, ми повинні пам'ятати про головну мету сучасних іспитів з ІМ і розглядати іноземну мову як засіб міжкультурного спілкування, про що провідний британський тестолог Чарлз Олдерсон висловився таким чином:

“Сучасні іспити не ставлять за мету з'ясування того, чи спроможні учні/студенти перекладати, перефразувати ізольовані слова або речення поза контекстом їхнього застосування. Метою сучасних іспитів є встановлення спроможності учнів/студентів спілкуватися іноземною мовою для досягнення взаєморозуміння в повсякденних ситуаціях у ролі мовця або слухача, в усній чи письмовій формі комунікації навіть на найнижчому рівні володіння мовою та без 100%-ої коректності оформлення висловлювань”².

О. Квасова (Київ), Л. Гнаповська (Суми)

¹ Alderson, Ch. J., Clapham C., Wall D. Language Test Construction and Evaluation. – Cambridge: Cambridge University Press, 199. – p. 9.

² Alderson, J. C., Cseresznyes, M. Into Europe. Prepare for Modern English Exams. Reading and Use of English. – Budapest, 2005. – p.13.

