
ПРАКТИЧЕСКА ФАЗА ОТ РАЗРАБОТВАНЕТО НА НОВ ТЕСТ

(анализ на резултатите и редакция
на “Тест по български като чужд, ниво А1”)

Галина Куртева, Десислава Димитрова
Нов български университет, София

This article describes the process of pre-testing a new test of Bulgarian as a foreign language which is conducted in several stages. Proper pre-testing requires a clear understanding of each task and question's intent and is used to identify and correct tools and instruments in test design. The authors rely not only on the literature on that subject, but also on the knowledge gained by practice, experience and collected evidence for test qualities. For the investigation of test qualities two kinds of statistical procedures are run. The feedback questionnaire for teachers' opinion and instruction for test administration was developed. After analyzing the conclusions are summarized. As a result some important decisions are made. They concern some changes in test specification: response format, rubrics, visual input, timing and weighting.

Разработването на нов изпит преминава през три последователни етапа: концептуална фаза, теоретична фаза и практическа фаза. Задачата на първите два етапа беше изработването на модел на теста. Теоретичната обосновка на този модел беше предмет на описание в друга статия¹. Тук детайлно са разгледани елементите на практическата фаза. Целта на този етап беше да се съберат данни за качествата на теста като инструмент за измерване и да се интерпретират резултатите по отношение на езиковата компетентност по български език като чужд, разбираана според скалите и дескрипторите на Общата европейска езикова рамка. Обработката и анализите на данните доведоха до експертни решения за промяна в модела на теста.

Работата по теста беше изцяло съобразена със стандартите за добра практика на АЛТЕ² и с основните теоретични принципи за валидност на измерването. Стандартите за добра изпитна практика се отнасят до всеки етап от изпитния процес. Те са групирани в няколко направления:

1. Стандарти за разработването на тест.
2. Стандарти за администрирането на изпит.

3. Стандарти за точкуването и оценяването.
4. Стандарти за анализ на данните.
5. Стандарти за комуникация със заинтересованите страни.

Стандарти за разработване на тест

Първата група стандарти се отнасят към концептуалната фаза на изпита. Те изискват теоретичната обосновка на изпита да се основава на разбирането за комуникативна езикова компетентност. При теоретичното конструиране на теста по български език за чужденци беше използвана Европейската езикова рамка за дефиниране на комуникативната езикова компетентност и нейните елементи³. Теоретичната връзка между използваните скали и дескриптори на рамката, от една страна, и задачите на теста – от друга, е описана в спецификацията на теста.

По отношение на теоретичната обосновка на изпита стандартите също така изискват да се проследи контекстуалната валидност на задачите. Това означава да се провери дали задачите създава условия за изява на езиковата компетентност, която искаме да измерим. Необходимо е да се постигне и теоретична валидност на измерването умение по отношение на неговата автентичност, т. е. дали измерването умение е наистина процес за извършване на говорене, слушане, писане и четене.

Задачата е сложен комплекс от инструкции, въпроси, материал за обработка, формат на отговора, условия за изпълнение и критерии за оценяване. За да бъде наблюдавано езиково изпълнение по всички елементи на езиковата компетентност, която е в основата на теоретичния модел, бяха определени 65 въпроса, групирани в четири раздела, и 9 задачи⁴. Всички инструкции са на български език, формулирани директно и с възможно най-малко думи, например: “Слушайте записа и изберете”, “Прочетете двете писма и отговорете на въпросите”. За да се избегнат евентуални езикови затруднения при разбиране на инструкциите, към всяка задача има пример за отговор. Въпросите, материалът за обработка и формата за отговор са взаимосвързани. Те предпоставят когнитивното натоварване при изпълнението на задачата и по този начин влияят на трудността. В първоначалния модел на теста тази трудност беше предвидена от експертите и въз основа на техния дългогодишен опит в преподаването на български като чужд език.

Въпросите са формулирани възможно най-ясно, за да бъдат лесно разбрани и да се избегне нежелана вариативност на отговора. Изборът на подходящ формат за отговора беше съобразен с вида на умението, което се проверява. Самият материал за обработка също има характеристики, които внимателно бяха подбрани, така че да е с подходяща дължина и подходящ в лингвистично отношение. В теста са използвани и визуални материали за обработка, оформени като цветни картини, за да са ясно и недвусмислено различими.

В задачите от раздел *Слушане* отговорът е избираем от 3 - 4 алтернативи, изразени невербално или със съвсем кратки думи “да”, “не” и надписи с дните от седмицата. Материалът за обработка е разнообразен за трите задачи: едно изречение,

диалог и съобщение (прогноза за времето). Той се предоставя слухово, с паузи и повторения. Инструкциите са оформени визуално в началото на всяка задача, но се изговарят и в материала за обработка.

В задачите от раздел *Четене* всички отговори са избираеми, без да изискват писане, а само отбелязване на избрания вариант. Материалите за обработка са кратки описателни текстове с разказни изречения и четири диалога с различна комуникативна цел: запознанство, молба, покана и т. н.

В задачите от раздел *Писане* се изисква отговор с една дума и с едно изречение. Материалът за обработка е съответно текст от писма, обяви и визуални опори, които насочват към създаването на изречение.

В задачите от раздел *Говорене* се изисква отговор с едно изречение. Материалът за обработка са въпроси за представяне и визуална опора за описание.

За да бъдат разграничени уменията като *слушане, четене, писане и говорене*, трябваше да се разработят предварително различни варианти на отделните задачи, така че да се намери най-добрата комбинация между начина на представяне на материала за обработка: визуален или слухов, и начина на отговаряне: с дума, с изречение, само с отбелязване и т. н.

Условие за изпълнение на отделните задачи не беше посочено. Беше предположено време за изпълнение на целия тест, което не беше известно на изпитваните, а само на преподавателите, провеждащи изпита. По този начин искахме да проверим колко време ще е достатъчно за изпълнение на теста.

Стандарти за администриране на изпит

Следващата група стандарти се отнасят към администрирането на изпита. Те изискват да има контрол върху условията за провеждане, така че да се гарантират необходимите според спецификацията на теста условия за провеждане на изпита, както и тяхната еднаквост за всички изпитвани лица. За тази цел беше разработена инструкция към преподавателите. В по-голямата част от случаите екипът, който разработи теста, администрираше изпита. По този начин бяха избегнати нежелани фактори като намеса от страна на преподавателите и бяха наблюдавани пряко някои важни условия за изпълнението: време на теста, яснота на инструкциите, яснота на визуалните опори, реакция на изпитваните и др.

Стандарти за точкуване и оценяване

Третата група стандарти се отнасят към точкуването и оценяването. Те изискват да има предварително определени критерии за оценяване, както и процедури за получаването на точки. В теста по български език за чужденци всяка задача оценява отделен аспект от лингвистичната и функционалната компетентност. Изявата на този аспект е съобразена със скалите от езиковата рамка. Те представляват критерият за изява на всеки отделен аспект. В теста са използвани скалите за:

- *Речников обхват*: разполага с елементарен набор от отделни думи и изрази, свързани с конкретни ситуации.

- *Обхват на езиковите средства*: разполага с елементарен набор от прости изра-

зи за предаване на информация, свързана с него и с непосредствените му нужди.

- *Граматична точност*: показва ограничено владение на някои прости граматични структури и готови изречения.

- *Владение на произношение и интонация*: произношение на много ограничен набор от запаметени думи и изрази, може да се разбере с известно усилие от носител на езика, привикнал към специфичния говор на езиковата група на говорещия.

- *Владение на правопис*: може да преписва кратки изрази и познати думи, названията на предмети от всекидневието, имена на магазини и някои често употребявани изрази. Може да произнесе буква по буква адреса, националността си и други лични данни.

- *Свързаност и последователност*: може да свързва съчетания от думи с елементарни съюзи като "и".

- *Свобода на изказа*: може да борави с кратки, отделни обикновени стереотипни фрази, като прави чести паузи, за да намери подходящите изрази и да произнесе непознатите думи или да възстанови прекъснатото общуване.

Посочените аспекти трябва да се отразят в изпълнението на задачата, което е и доказателство за верността на отговора.

Процедурите по измерването включват два последователни етапа: определяне на брой критерии за вяръност и процедури за поставяне на точка. В седем от задачите има един критерий за вяръност, което означава, че всеки верен отговор носи една точка. Резултатът се натрупва чрез броене на верните отговори. В две от задачите точките се присъждат според предварително определен критерий и оценяващият преценява дали този критерий се проявява в изпълнението на студентите. Това са задачите за писане на изречение и за говорене. За оценяване на писането и говоренето бяха формулирани по три критерия за вяръност, като оценяващият определя наличието или отсъствието им.

Критерии за оценяване на писането:

- Пълнота на изречението: има подлог, сказуемо, допълнение.

- Граматична правилност на изречението: съгласуването е вярно.

- Граматична правилност на думата: членуване, род, число.

Критерии за оценяване на говоренето:

- Комуникира успешно: разбрал е въпроса и дава някакъв отговор.

- Комуникира правилно: отговаря в изречение с подходящите думи.

- Комуникира разбираемо: произнася думите с акцент, но вярно.

Авторите на теста бяха и оценяващи. Всяка работа беше оценена според нейния суров бал, без да бъдат поставяни оценки по шестобалната скала.

Стандарти за анализ на данни

Четвъртата група стандарти се отнасят към анализ на данните. Те изискват да се приложат статистически процедури както за описание на тестовия резултат, така и за трудността на всеки отделен въпрос. При анализа бяха приложени процедури от описателната статистика и IRT⁵ анализи, с помощта на специално разработен софтуер⁶.

Последната група стандарти, които се отнасят към комуникацията със заинтересованите страни, изисква да се предоставят възможности за лесна и достъпна връзка с изпитваните. За тази цел се предвижда създаването на сайт. Тъй като тук се описва фазата на апробиране⁷ на теста, тази група стандарти няма да бъде разглеждана.

Апробирането се извърши в продължение на пет месеца. Нашата цел беше да проверим дали очакваните качества на теста ще бъдат подкрепени от реалните данни. Събраната информация се отнася към всяка група от горепосочените стандарти, които сме използвали като критерии за всеки етап от изпитния процес.

Данни за изпълнението на теста бяха събрани от 214 студенти от различни националности, предимно във възрастовата граница между 18 и 21 години. Те се обучават в 10 университета, 8 в страната⁸ и 2 извън страната. Обучението е в курсове, самостоятелно или частно. Таблица 1 дава информация за броя и националната принадлежност на студентите, които са направили теста.

Страна	Брой на отделните случаи	Процент спрямо общия брой участници
Германия	13	6,1
Гърция	23	10,7
Испания	3	1,4
Италия	1	,5
Казахстан	1	,5
Кипър	9	4,2
Китай	5	2,3
Куба	2	,9
Либия	1	,5
Молдова	2	,9
Нигерия	1	,5
Палестина	1	,5
Полша	2	,9
Португалия	1	,5
Русия	2	,9
САЩ	3	1,4
Турция	139	65,0
Франция	1	,5
Холандия	1	,5
Чехия	2	,9
ЮАР	1	,5

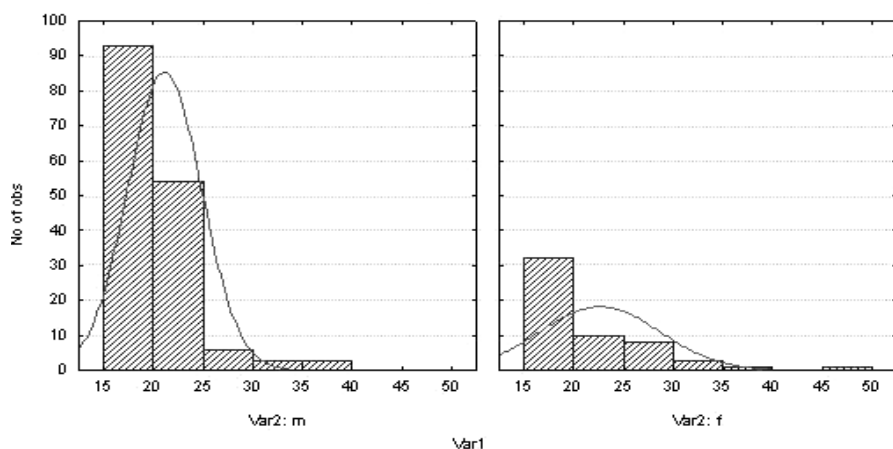
Таблица 1. Брой и национална принадлежност на студентите, направили теста

Откроява се големият брой на студенти от Турция – 139. Това е закономерно, като се има предвид, че в момента турските студенти съставляват по-голяма част от чуждестранните студенти в България.

По същата причина и мъжете са повече от жените, както се вижда от фигура 1. И за двата пола най-многобройни са участниците във възрастовата група между 18–20 години.

m = мъже, f = жени

Фигура 1. Представяне на участниците според пола и възрастта⁹



Всички студенти изпълняват един и същ вариант на теста. По време на изпълнението студентите имаха право да задават въпроси за задачите: ако не са разбрали инструкциите, ако не могат да се ориентират какъв е форматът на отговор, ако не разбират дума от текста. По време на апробирането, освен изпитните материали от студентите, бяха документирани и всички останали впечатления за провеждането на теста – въпросите на студентите, времето, за което изпълняват теста, наблюдения на преподавателите, неочаквано възникнали проблеми. Тази информация също се оказва много ценна при решенията за промяна в модела на теста.

Резултатът от всеки раздел беше анализиран поотделно.

При раздел *Слушане* верните отговори се натрупват чрез сумиране и максималният резултат е 11 точки (вж. таблица 2).

Брой студенти	Минимален бал	Максимален бал	Средна стойност	Стандартно отклонение
214	4	11	10,06	1,2

Таблица 2. Описание на суровия бал от раздел *Слушане*

Раздел *Слушане* обхваща три задачи, които са степенувани по трудност. Текстовете са предварително записани и не се четат от преподавател или квестор, а се слушат два пъти. В първата задача изпитваните слушат кратки изречения и трябва да отбележат правилния отговор. Във втората задача те слушат кратък диалог с визуална опора и също трябва да изберат правилната картинка. Третата задача е най-дълга, всъщност това е прогноза за времето за три дни. Студентите имат пред себе си символни изображения на времето за четири дни, от които трябва да изберат три.

Таблица 3 представя резултатите от трите задачи на раздел *Слушане*.

Бал	Честота	Процент
10	74	34,6
11	96	44,9
4	2	,9
6	2	,9
7	6	2,8
8	12	5,6
9	22	10,3
Общо:	214	100,0

Таблица 3. Резултат от раздел *Слушане*

Както се вижда от резултатите, 96 студенти имат максимален резултат, а 74 студенти имат само по един грешен отговор. IRT анализите показаха, че този грешен отговор е допускан най-често във въпросите на третата задача от раздел *Слушане*. Причина за затруднението е формата на отговор, а не знанието на студентите. Там се получи объркване от наличието на четири изображения при прогноза на времето за три дни. Много от изпитваните не разбраха добре задачата и това наложи намеса на квесторите, както и допълнителни разяснения по задачата.

При анализа и корекциите на теста след апробацията бе решено третата задача да бъде променена поради липса на достатъчно яснота и защото затрудни много голяма част от студентите. Бяха оставени три изображения, които да бъдат адекватни на трите дни, отразени в прогнозата.

В раздел *Четене* верните отговори се натрупват отново чрез сумиране на ве-

рните отговори. В раздела има повече въпроси и затова максималният бал е 32 (вж. таблица 4).

Брой студенти	Минимален бал	Максимален бал	Средна стойност	Стандартно отклонение
214	11	32	27,78	4,1

Таблица 4. Описание на суровия бал от раздел *Четене*

Разделът *Четене* съдържа също три задачи с различна степен на трудност. В първите две задачи е необходимо да се попълнят празните места в кратък свързан текст. Студентите трябва да изберат една от двете предложени опции. Текстовете са с позната тематика от ежедневието и рутинни дейности и имат за цел да се проверят граматичните знания на изпитваните. Въпреки че тези две задачи не са сложни и голяма част от изпитваните са се справили с тях, се наблюдават и значителен брой грешки. Те не се дължат на неразбиране (студентите не задаваха въпроси и споделиха, че този раздел е ясен и точен), а на недостатъчно познаване на граматичните правила. Тази част от раздела има не толкова комуникативна, колкото граматична насоченост. Изреченията и задачите са свързани с конкретни граматични категории и задачите имат за цел да измерят както определени граматични знания, така и тяхното правилно прилагане. Тук нямаше неясноти при задачата, студентите са спазили изискванията, но бяха допуснати редица грешки, въпреки че изборът беше само между две опции.

Третата задача от този раздел е комуникативна и съдържа различни ситуации. Състои се от четири ситуативни диалога с предложени към тях изречения. Необходимо е да се попълнят липсващите реплики, като са предложени с една опция повече от необходимите позиции. За повече яснота и прегледност задачата съдържа и пример. С нея се проверява разбирането и реакцията на студентите в различни ежедневни ситуации, комуникативната им способност, а също и уменията им да съставят кратък, свързан и логичен текст. И тук изпитваните са се справили добре с всички ситуации, с изключение на една. Въпреки че диалогът е пределно ясен, при него се наблюдават най-много неверни отговори. Това наложи след апробацията да се направят корекции от авторите на теста. Промените целят да се улеснят разбирането и правилният избор от страна на студентите.

От таблица 5 се вижда, че в изпълнението на студентите има по-голямо разнообразие на бала, което означава, че в раздел *Четене* задачите и въпросите разграничават добре знаещите от незнаещите студенти. Това се потвърди и от IRT анализите за трудността на всеки отделен въпрос.

Бал	Честота	Процент
11	1	,5

12	1	,5
14	1	,5
17	3	1,4
18	2	,9
19	3	1,4
20	3	1,4
21	7	3,3
22	6	2,8
23	6	2,8
24	6	2,8
25	10	4,7
26	8	3,7
27	16	7,5
28	22	10,3
29	24	11,2
30	32	15,0
31	35	16,4
32	28	13,1
Общо	214	100,0

Таблица 5. Резултат от раздел *Четене*

Раздел *Писане* включва инструменти за проверка както на продуктивното, така и на репродуктивното писане. Състои се също от три степенувани по трудност задачи, с различен максимален брой точки.

Първата задача представлява две свързани кратки електронни писма. Двата имейла са свързани, като вторият представлява отговор на първия имейл. Изпитваният трябва да отговори на въпроси върху двата текста. Форматът на отговора изисква да се отговори с дума или с изречение, без това да е уточнено в инструкцията на задачата. Това създаде определена трудност при оценяването на отговорите, тъй като имаше разнообразни отговори: от отделни думи до цели изречения. С оглед на спецификацията и критериите за вярност бяха приети за верни всички отговори, които бяха точни по отношение на зададения въпрос, независимо дали съдържат една дума, фраза или цяло изречение. Бе решено да бъдат направени редакции в инструкцията на задачата.

Втората задача обхваща три кратки текста – обяви за работното време на различ-

ни обекти. Изисква се отговор с една дума. В изпълнението на тази задача нямаше трудности и тук общият бал точки е доста висок.

Таблица 6 показва максималния бал за двете задачи от раздел *Писане*, при които общият бал се натрупва чрез сумиране на верните отговори.

Брой студенти	Минимален бал	Максимален бал	Средна стойност	Стандартно отклонение
214	0	8	6,44	1,8

Таблица 6. Описание на суровия бал от раздел *Писане*, репродуктивни задачи

Резултатът от двете задачи (вж. таблица 7) показва, че 81 студенти имат максимален бал за тази задача. Грешните отговори са предимно при въпросите за разбиране на съдържанието на двете писма. Въпросите от тази задача се оказват особено подходящи за измерване на началната езикова компетентност. Те разпределят добре студентите според трите заложи в спецификацията критерия за изявата им: свобода на изказа, владение на правопис и граматична точност.

Бал	Честота	Процент
0	4	1,9
1	1	,5
2	3	1,4
3	14	6,5
4	12	5,6
5	12	5,6
6	36	16,8
7	51	23,8
8	81	37,9
Общо:	214	100,0

Таблица 7. Резултат от раздел *Писане*, репродуктивни задачи

Третата задача от раздел *Писане* изисква продуктивно писане и е с най-висока степен на трудност. Тя се базира на шест визуални опори. Студентите трябва да напишат по едно изречение, описващо изображението на картината. Тематиката е върху рутинни ежедневни дейности. Всяко изречение беше оценено по три критерия: пълнота на изречението, граматична правилност на изречението и граматична правилност на думата. Точките се натрупват след преценка от страна на оценяващия дали съответен критерий отговаря на дадения отговор.

Максималният бал за тази задача е 18. Той се получава при шест съставени вярно и правилно изречения (вж. таблица 8).

Брой студенти	Минимален бал	Максимален бал	Средна стойност	Стандартно отклонение
214	0	18	11,65	5,2

Таблица 8. Описание на суровия бал от раздел *Писане*, продуктивна задача

Разпределението на тестовия бал може да се види от таблица 9. Изобщо не са отговорили на задачата 17 студенти. Приблизително толкова са и студентите, които са получили максимален бал – 20. Тук бяха регистрирани най-много грешки – по отношение на граматиката, съгласуването, свързаността и верността на текста. Разнообразието от тестовия бал показва, че тази задача дава много информация за разликите в началната езикова компетентност. Затова независимо от нейната относителна трудност авторският екип реши да запази тази задача в същия формат.

Бал	Честота	Процент
0	17	7,9
1	2	,9
3	2	,9
4	3	1,4
5	6	2,8
6	5	2,3
7	13	6,1
8	7	3,3
9	12	5,6
10	6	2,8
11	14	6,5
12	11	5,1
13	15	7,0
14	20	9,3
15	19	8,9
16	22	10,3
17	20	9,3
18	20	9,3
Общо:	214	100,0

Таблица 9. Резултат от раздел *Писане*, продуктивна задача

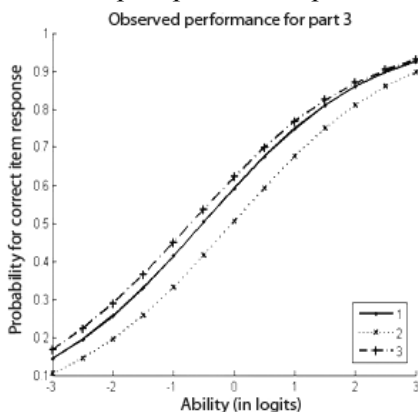
Трудността на всеки критерий спрямо всяко едно от шестте изречения беше проверена чрез IRT анализ. В случая бе важно дали тези критерии са достатъчно основание да се разграничат студентите според тяхната начална езикова компетентност и дали всеки критерий дава представа за това, т. е. дали всеки критерий

работи. Фигура 2 дава пример за такъв анализ.

По вертикалната ос се отбелязва вероятността за верен отговор, която се отчита от 0 до 1. По хоризонталната ос се отбелязва степента на умението, в случая езиковото умение най-общо, по скала със среда 0 и стандартно отклонение 1. От примера се вижда, че кривите, които представляват трите критерия поотделно, започват от много ниски стойности. Това показва, че за студенти с много ниска степен на уменията вероятността да преодолеят критерия е много малка. Трудността на критериите се изчислява в стойност от (-2) до $(+2)$. За критерий номер 2 тя е най-висока $(-0,02)$, за критерий номер 3 е най-ниска $(-0,7)$, критерий номер 1 има стойност $(-0,5)$. Във всички изречения критерий номер 2 (граматична правилност на изречението) се оказва най-труден за преодоляване.

Дискриминативната сила на критериите се определя от извивката на кривата. Макар че този параметър може да приема стойности от минус безкрайност до плюс безкрайност, на практика се наблюдават стойности от 0 до 2. В примера от фигура 2 всички критерии имат почти еднаква дискриминативна сила. В случаите с останалите изречения дискриминативната сила се различава повече за трите критерия, но във всички случаи те работят при всяко изречение. Като заключение можем да кажем, че критериите са успешно формулирани и на този етап няма да бъдат променени.

Фигура 2. Трудност на критериите за изречение номер 3¹⁰



Разделът *Говорене* е последният и проверява произношението, комуникативните и говорните умения на изпитваните. Състои се от две части:

- Представяне – студентът трябва да се представи по име, възраст, години, занятие в момента и т. н. (той получава въпроси, на които отговаря).

- Коментар на визуална опора – дава му се картина или фотография, която той трябва да коментира (под нея има няколко подпомагащи въпроса като: Какво виждате на картината? Какво прави този човек? Къде е той? Какъв е той според Вас? и др.).

Въпросите за представянето бяха едни и същи за всички студенти, а визуалните опори се раздаваха на случаен принцип. Отговорът на всеки въпрос беше оценен по три критерия: комуникира успешно (дава знак, че е разбрал въпроса и дава някакъв отговор); комуникира правилно (отговаря в изречение с подходящи думи), комуникира разбираемо (произнася думите с акцент, но вярно, така че да се разбира смисълът). На таблица 10 се вижда, че максималният бал за този раздел е 24 точки, които се получават от отговор на 8 въпроса.

Брой студенти	Минимален бал	Максимален бал	Средна стойност	Стандартно отклонение
214	2	24	17,25	6,2

Таблица 10. Описание на суровия бал от раздел *Говорене*

Поради разнообразието от въпроси максималният бал е достигнат само от 54 студенти. Всички останали са изпълнили само някои от критериите. Таблица 11 показва разпределението на тестовия бал от раздел *Говорене*. Той достатъчно точно разграничава студентите според тяхното говорно умение.

Бал	Честота	Процент
2	1	,5
3	3	1,4
6	8	3,7
7	7	3,3
8	9	4,2
9	8	3,7
10	3	1,4
11	10	4,7
12	7	3,3
13	7	3,3
14	8	3,7
15	8	3,7
16	8	3,7
17	7	3,3
18	10	4,7
19	13	6,1

20	12	5,6
21	11	5,1
22	8	3,7
23	12	5,6
24	54	25,2

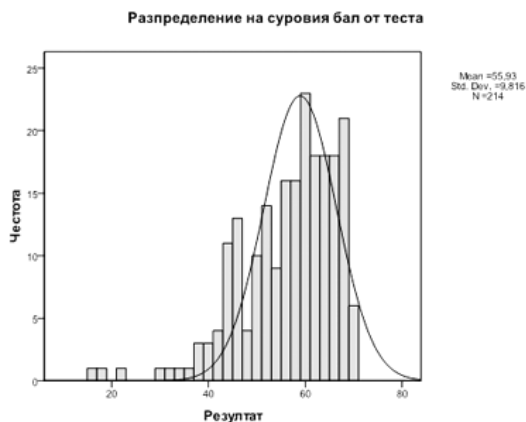
Таблица 11. Резултат от раздел *Говорене*

Въпреки че въпросите от този раздел не са особено трудни и тематиката е добре позната на всички изпитвани, те се затрудниха много при раздела *Говорене*, което налагаше честата намеса на квесторите с насочващи и подпомагащи въпроси. Анализът на тази част показва, че това се дължи както на липсата на достатъчно добре развити говорни навици, така и на психологически прегради, породени от присъствието на непознати хора, от ограничението във времето, а също и от наличието на записваща техника, която особено смущаваше изпитваните.

Трудността на критериите за говорене беше изследвана чрез IRT анализ, така както и критериите за писане на изречения. Данните показват, че отново критерий номер 2 се оказва най-труден за преодоляване (комуникира правилно: отговаря в изречение с подходящи думи). Дискриминативната сила и на трите критерия е в приемливи стойности, така че всеки критерий да дава информация за уменията на студентите. Това дава основание да се приеме, че критериите са подходящо формулирани.

Резултатът от всички части на теста е описан на фигура 3. Той показва, че голяма част от студентите са се справили успешно. Това е приемливо за предназначението на изпита по български език за равнище A1. Екипът е заложил очакване, че резултатите от изпита ще се използват като обратна връзка за преподавателите по отношение на образователния дизайн на курса и обратна връзка за самите студенти относно силните и слабите страни от обучението им по български език. Затова за нас границата за успешно преминаване ще е само една. Доказателство за умение на равнище A1 за владеење на български като чужд език ще е натрупването на 50% верни отговори. По отношение на обратната връзка към самите студенти екипът ще насочи своите усилия към измерване с цел когнитивно оценяване, което е заложено като принцип при разработването на теста, но ще е предмет на бъдещи изследвания.

Като цяло апробирането на теста показва, от една страна, пропуските и трудностите в отделните тестови задачи, а от друга – даде ясна и пълна картина за нивото и характеристиките на изпитваните. Тук бяха отчетени националността, възрастта, мястото и степента на изучаване на български език, целите на изучаване на езика, и др.



Фигура 3. Описание на резултата от всички раздели на теста.

Като следствие от обсъждането на промени в модела бяха направени уточнения на спецификацията по отношение на избора на текстове за обработка, уточнения на спецификацията по отношение на детайлите на визуалните опори, уточнения на спецификацията за формата на отговора при някои задачи, нови инструкции към някои задачи, уточнение на времето за изпълнение на теста, промяна в тежестта на резултата от отделните раздели, така че да се балансират отделните умения: слушане, четене, писане и говорене.

Предстои още сериозна работа по обработката и детайлното анализиране на резултатите, по прилагане на критериите, уточняване на изискванията и стандартите, и др. Изпробването на теста в реална среда и при автентична ситуация налага извършване на корекции и промяна на тези задачи и позиции, които показаха отклонения и дефект при претестването. Аprobацията и анализът на резултатите дадоха възможност да се съберат ценни данни и информация, които могат да бъдат използвани в бъдеще както при създаването на нови тестове за различни цели, така и при съставянето на учебници, помагала и електронни материали при обучението и оценяването на българския език като чужд.

БЕЛЕЖКИ

¹Дизайн и разработване на нов тест (Български език като чужд – ниво А1) ”

²Асоциация на европейските центри за изпитване по чужд език

³Теоретичната обосновка е описана в “Дизайн и разработване на нов тест (Български език като чужд - ниво А1) ”

⁴Раздел наричаме обособяването на част “Слушане”, “Четене”, “Писане” и “Говорене” в теста. Задача наричаме компонент от теста, който съдържа въпроси, насочени от една и съща инструкция.

⁵Item response theory, теория за отговор на тестови въпрос

⁶Автор на софтуерния продукт е д-р Димитър Атанасов, ст. ас. в Департамент по информатика, НБУ.

⁷Има се предвид практическата фаза от разработването на нов тест

⁸София: Медицинска академия, Национална спортна академия, Технически университет, Нов български университет, Пловдив: Аграрен университет, Университет по хранителни технологии, Русе: Русенски университет, САЩ: Университет в Питсбърг, Германия: Университет “Хумболт”.

⁹Данните са обработени от Димитър Атанасов

¹⁰Данните са обработени от Димитър Атанасов.

ЛИТЕРАТУРА

ALTE: ALTE Code of practice. Available from: <http://www.alte.org>

ALTE: Minimum Standards for establishing quality profiles in ALTE examinations. Available from: <http://www.alte.org>

ALTE: Quality Management Checklists. Available from: <http://www.alte.org>

Bachman, L. (2004): *Statistical Analyses for Language Assessment*. Cambridge University Press.

Weir, C. (2005): *Language Testing and Validation*. New York: Palgrave Macmillan.