



НОВ БЪЛГАРСКИ УНИВЕРСИТЕТ
МАГИСТЪРСКИ ФАКУЛТЕТ
ДЕПАРТАМЕНТ „РОМАНИСТИКА И ГЕРМАНИСТИКА”
ДП „Методика на обучението по съвременни езици“

Димитрина Иванова Гергова

ОЦЕНЯВАНЕТО ПО ЧУЖДИ ЕЗИЦИ
В БЪЛГАРСКОТО УЧИЛИЩЕ

АВТОРЕФЕРАТ

на дисертационен труд за присъждане
на образователна и научна степен „доктор”

Област на висше образование: 1. Педагогически науки
Професионално направление: 1.3. Педагогика на обучението по ...
Научна специалност: 05.07.03. Методика на обучението по немски език

Научен консултант: проф. д-н Павлина Стефанова

Рецензенти:

доц. д-р Светлана Димитрова-Гюзелева

доц. д-р Нели Митева

София
2016

Дисертационният труд е в обем от 190 страници, от които 155 страници основен текст и 12 страници библиография. Състои се от: увод, три глави, изводи, заключение, приноси, използвана литература и приложения. В текстовата част са включени 29 таблици и 14 фигури. Библиографията включва 120 заглавия, от които 31 – на български език, 83 – на немски език, 6 – на английски език и 11 – интернет източници. Приложенията са 7 с общ обем 21 страници.

Дисертацията е осъществена под научното ръководство на проф. д-н Павлина Стефанова.

Дисертационният труд е обсъден и насочен за защита на заседание на разширения съвет на департамента „Романистика и германистика“ при МФ на НБУ, проведено на 18 юли 2016 г.

Защитата на дисертационния труд ще се състои на 31.10.2016 г. от 11.20 часа в зала ... на НБУ. Научното жури, определено със Заповед № З-РК-423 на Ректора на НБУ, е в следния състав:

Вътрешни членове: 1. доц. д-р Светлана Димитрова-Гюзелева (НБУ, председател)

2. проф. д-н Павлина Стефанова (НБУ)

Външни членове: 3. проф. д-р Даниела Стойчева (СУ „Св. Климент Охридски“)

4. доц. д-р Йорданка Симеонова (СУ „Св. Климент Охридски“)

5. доц. д-р Нели Митева (ШУ „Епископ Константин Преславски“)

Съдържание

Увод	3
Глава 1. Обща проблематика на измерването на постиженията в езиковото обучение	
1.1. Теория на теста	9
1.2. Видове оценяване и категоризиране на тестовете	11
1.3. Функции на оценяването на постиженията в езиковото обучение	14
1.4. Критерии за качество на езиковите тестове и тестовите задачи	16
1.5. Етапи при създаването на тестовете и видове тестови формати	17
Глава 2. Методи за разработване и оценка на тестови задачи за оценяване на езикови компетентности	
2.1. Принципи за създаване на комуникативни езикови тестове	18
2.2. Оценяване на уменията <i>четене с разбиране</i>	20
2.3. Връзката между вида на текста, вида на задачата и целта на четене	22
2.4. Текстове за тестови задачи за оценяване на уменията <i>четене с разбиране</i>	23
2.5. Правила за създаване на тестови задачи с избираем отговор и критерии за тяхната експертна оценка	24
2.6. Анализ на тестове	27
Глава 3. Експериментална проверка на качествата на тестови формати с избираем отговор	
3.1. Представяне на процеса на разработване на теста за експеримента	29
3.2. Представяне на проведения експеримент: изготвяне на извадка, апробация на теста	30
3.3. Статистическа обработка и анализ на резултатите от теста	30
3.4. Изводи от експерименталната проверка	37
3.5. Представяне на анкетата и анализ на резултатите	38
Заклучение	40
Приноси	42
Библиография	43
Списък с публикациите, свързани с темата на дисертацията	49
Декларация за авторство	50

Увод

Важното значение на оценката на постиженията за развитието и шансовете за успех в живота на всеки ученик определя практиката на оценяването като централна тема при обсъждане на проблемите на образованието. Изпитването, тестването и оценяването по езици си поставя като основна цел проверката на езиковите компетентности. Преди да се пристъпи към оценяване, е важно да се определи обектът на проверка. В езиковото обучение се приема, че обект на проверка са компонентите на комуникативната компетентност, като нейното измерване се ориентира към критерии, произтичащи от целите и съдържанието на съответното езиково ниво или от предварително определените цели за съответния етап на обучение. Когато в обучението се използват готови тестове, трябва точно да се знае какви компетентности се проверяват чрез тези тестове, за да е целесъобразна употребата им. Когато учителите разработват сами тестове, трябва първо да уточнят какво точно искат да проверят и как може да се направи това в съответствие с критериите за качество на езиковите тестове. Необходимо е също така да се отговори на въпроса дали изводите, които ще бъдат направени въз основа на тестовете/ изпитите, и дали решенията, които ще бъдат взети, ще са справедливи за учениците. Затова формирането на диагностични компетентности е важна част от професионалната подготовка на учителите и важна предпоставка за провеждане на добро обучение.

Традиции в стандартизираното оценяване по чужди езици в България се създават през последните десетилетия. В програмите за развитие на образованието в България, приети през 2006 година¹ и през 2009 година², основен приоритет се дава на ориентирането на българското училищно образование към изграждане на компетентности и формиране на практически умения. Като едно от най-важните условия за постигане на европейско качество на образованието са определени изграждането и усъвършенстването на системата за външно и вътрешно оценяване. Целите на националното (външното) оценяване са свързани с проследяване и диагностициране на развитието на знанията на учениците по време на обучението им, установяване на нивото на постигане на държавните образователни изисквания (ДОИ) за съответния етап или степен и използване на резултатите като селекция за следваща

¹Национална програма за развитие на училищното образование и предучилищното възпитание и подготовка (2006 – 2015)

²Програма за развитие на образованието, науката и младежките политики в Република България (2009-2013)

образователна степен. Утвърдената система за национално оценяване позволява обективна оценка на постигнатите резултати в образователния процес. Разработеният инструментариум (дидактически тестове) от ЦКОКУО, предоставен на МОН за провеждане на национални оценявания, гарантира надеждността и валидността на оценяването. В годините на реформи българската образователна система е изправена пред редица предизвикателства, свързани с усъвършенстване на методите на преподаване и на изпитните материали за вътрешно и външно оценяване.

Нерешените въпроси в практиката за измерване на постиженията на учениците по чужди езици налагат необходимостта от нови изследвания в областта на оценяването в българското училище. Това е първият мотив, който определи избора на темата на настоящия дисертационен труд. Другият важен мотив е същественото значение на осигуряването на ефективно обучение по езици, отговарящо на действителните потребности на учениците, и на адекватни методи и инструменти за оценяване на езиковите компетентности.

Резултатите на българските ученици от международните оценявания PISA³ за четивна грамотност и от ESLC⁴ – за езикови компетентности са по-ниски от средните резултати на учениците от държавите, взели участие в изследванията. Тези данни са индикатор, че образователната ни система е ориентирана предимно към осигуряване на теоретично познание, а не на практически умения, което е в разрез със съвременните концепции за образование и с реалните потребности на учениците. За да се подобри балансът между теоретичната и практическата подготовка, е необходимо преосмислянето както на методите на преподаване, така и на методите на оценка на постиженията. Изборът на темата е свързан с важността на изследванията на оценяването на езиковите компетентности за осигуряване на качеството на изпитните материали (езиковите тестове), за обучението и квалификацията на преподавателите по езици, на авторите, рецензентите и оценителите на тестови задачи. В този аспект настоящият дисертационен труд се фокусира върху изследването на теоретичните основи на тестологията, същността и функциите на оценяването, критериите за качество на езиковите тестове и тестовите формати в комуникативните тестове с техните предимства и недостатъци. В дисертационния труд се акцентира върху методите за разработване и критериите за оценка на тестови задачи за оценяване на

³Виж: Петрова, С. (2010); Петрова, С. (2013); Петрова, С., Н. Василева (2013).

⁴Виж: Василева, Н. (2013)

езикови компетентности. Особено внимание е отделено на тестовите задачи, тъй като чрез тях се проверява наличието на определени умения и способности, съответно на постигането на компетентностите. Тестовите задачи дават ясен израз на абстрактно формулираните цели и очакваните резултати в учебните програми в конкретни изисквания. В рамките на изследването са разработени и подложени на проверка три различни тестови формати с избираем отговор за оценяване на уменията *четене с разбиране*, като са обхванати ученици от VIII клас в профилирани гимназии и профилирани паралелки на СОУ с интензивно изучаване на немски език. Чрез методите на класическата тестова теория са проверени качествата на тестовите задачи и тяхната пригодност за измерване на различните аспекти на уменията *четене с разбиране*. Анкетното проучване сред учениците, участвали в апробацията на тестовите задачи, предоставя информация за техните нагласи към оценяването на постиженията и мнението им за тестовите задачи, чрез които се оценяват техните езикови постижения.

Обект и предмет на изследването

Обект на настоящото изследване са тестовете (тестовите задачи) като инструмент за оценяване на комуникативните компетентности в обучението по съвременни езици като чужди езици.

Предмет на изследването са принципите и методите за разработване на валидни и надеждни тестове като адекватни инструменти за измерване на езикови компетентности. Експерименталната проверка на качествата (надеждност, трудност, дискриминативна сила) на различни тестови формати с избираем отговор се извършва в рамките на конструкта *четене с разбиране*.

Цели и задачи на изследването

Целите на дисертационния труд са чрез анализ на характеристиките и функциите на оценяването в езиковото обучение и принципите на комуникативните тестове да бъдат изведени и систематизирани правила за създаване на тестови задачи с избираем отговор и критерии за тяхната експертна оценка. Въз основа на анализ на данните от апробиране на тестови задачи да се изследват качествата на различни тестови формати с избираем отговор за измерване на уменията *четене с разбиране*. На базата на данни от анкетно проучване сред учениците, участвали в апробацията на

тестовите задачи, да се направят изводи за техните нагласи към оценяването и мнението им за тестовите задачи, чрез които се оценяват техните езикови постижения.

За постигане на целите изследването си поставя следните **задачи**:

- да се проучи и анализира достъпната научна литература и да се очертаят характеристиките, функциите и принципите на оценяването в чуждоезиковото обучение;
- да се анализират критериите за качество на езиковите тестове, характеристиките на тестовите формати и принципите за създаване на комуникативни тестове за оценяване на езикови умения и способности;
- да се систематизират правила за създаване на задачи с избираем отговор;
- да се разработят критерии за експертна оценка на текстове и тестови задачи за проверка на уменията *четене с разбиране*, предназначени за преподаватели, автори на тестови задачи и експерти;
- да се разработят тестови задачи за проверка на уменията *четене с разбиране* от различни формати с избираем отговор, като се спазят вече систематизираните правила;
- да се оценят тестовите задачи чрез разработените критерии за експертна оценка;
- да се апробират различни тестови формати с избираем отговор за оценка на уменията *четене с разбиране*;
- да се изследват психометричните характеристики на апробираните тестови задачи и да се направят изводи за техните качества;
- да се проучат нагласите на учениците към оценяването на постиженията и мнението им за тестовите задачи, чрез които се оценяват техните езикови постижения.

Хипотези на експерименталното изследване

На експериментална проверка се подлагат следните хипотези:

- Включването в структурата на тестовете на различни тестови формати с избираем отговор и на различни видове текстове, с които се проверяват различни стилове и стратегии на четене, дава възможност за проверка на различни аспекти на измерваното умение и за обосновани изводи относно проверяваните умения за *четене с разбиране*.

- От различни тестови формати с избираем отговор за проверка на умениято *четене с разбиране*, разработени чрез прилагане на правилата за създаване на задачи с избираем отговор и критериите за тяхната експертна оценка, могат да се съставят субтестове с еднаква трудност за ученици на едно и също езиково ниво.
- Тестването и оценяването е мотивиращ фактор за учащите и въздейства положително върху нагласите към ученето.

Методи на изследване

Изследователските методи са съобразени с характера на изследвания предмет, целите и задачите на изследването. Използвани са следните **методи**:

- анализ и синтез на теоретични схващания в достъпната научна литература по темата на изследването;
- анализ на структурата на тестовете за НВО по чужди езици;
- тест по немски език;
- анкета (за ученици);
- статистическа обработка на резултатите от апробацията на тестовите задачи и анализ на статистиките на задачите и на теста като цяло;
- статистическа обработка, анализ и интерпретация на резултатите от проведената анкета.

Структура и съдържание на дисертационния труд

Дисертационният труд се състои от увод, три основни глави, заключение, библиография и приложения.

В **увода** е обоснована актуалността на изследването, формулирани са обектът, предметът, целите и задачите, които трябва да бъдат изпълнени за тяхното постигане. Посочени са хипотезите на експерименталното изследване и методите на изследване.

В **първа глава** изследването се фокусира върху общите проблеми на измерването на постиженията в езиковото обучение. То е насочено към проучване на развитието на тестологията и в частност – на теорията на езиковите тестове. Специално внимание е отделено на мярката за измерване на езиковите постижения, видовете оценяване и видовете тестове. Синтезирано са представени функциите на оценяването на постиженията в езиковото обучение и критериите за качество на езиковите тестове,

като е отделено внимание на типичните грешки при тестването и оценяването. Представени са основните етапи при разработването на тестове и са разгледани характеристиките на тестовите формати в комуникативните тестове с техните предимства и недостатъци.

Във **втора глава** на дисертацията са изведени принципите за създаване на комуникативни тестове за оценяване на езикови умения, като фокусът е поставен върху оценяването на уменията *четене с разбиране*. Представени са критериите за избор на текстове за тестови задачи и детерминантите за оценяване на трудността на тестовете (текстове и тестови задачи) за оценяване на уменията *четене с разбиране*. От проучената литература са систематизирани правила за създаване на задачи с избираем отговор и са изведени критерии за експертна оценка на текстове и тестови задачи за проверка на уменията *четене с разбиране*. Направеният анализ на структурата на изпитните материали по чужди езици за външно оценяване в VII клас и VIII клас цели проверка на тяхното съответствие с наложените критерии за комуникативно ориентирани изпити от водещите европейски езикови организации. От проучената литература са направени изводи за нагласите към оценяването.

В **трета глава** на дисертацията е представена подготовката и провеждането на експерименталната проверка на качествата на тестови задачи с избираем отговор. Представените във втора глава методи за разработване и оценка на тестови задачи стават база за разработването на теста за апробация. Задачите за теста са разработени при спазване на систематизираните правила за създаване на тестови задачи с избираем отговор и са оценени чрез дефинираните критерии за експертна оценка на текстове и задачи с избираем отговор за проверка на уменията *четене с разбиране*. В апробацията на теста и попълването на анкетата участват 209 ученици, които учат немски език в VIII клас в профилирани гимназии и профилирани паралелки на СОУ с интензивно изучаване на чужд език. Представени са данните от статистическата обработка на резултатите от апробацията на тестовите задачи и от анкетата. Въз основа на анализа на статистиките на задачите и на теста като цяло са направени изводи за техните качества и тяхната пригодност за измерване на различните аспекти на уменията *четене с разбиране*. Анализът и интерпретацията на резултатите от проведената анкета даде възможност да бъдат направени изводи за нагласите на учениците към оценяването на постиженията и мнението им за тестовите задачи, чрез които се оценяват техните езикови постижения.

Глава 1.

Обща проблематика на измерването на постиженията в езиковото обучение

1.1. Теория на теста

В тази част от настоящата работа са представени тестологията, нейното възникване, развитие и приносът на редица изследователи за нейното развитие. Според проучената литература развитието на тестологията започва в края на XIX век с изследванията на Франсис Галтон, Джон Кейгъл и Дж. Фишер. Важна стъпка е направена в САЩ в началото на XX век от Дж. Райс, който тестира езиковото развитие на ученици и въз основа на резултатите от теста прави предложение за реформи в образованието. За развитието на теста като инструмент на научно изследване и разработването на методология за проверка на неговите качества (надеждност и валидност на резултатите) важен принос имат Карл Пирсон, Алфред Бине, Теодор Симон, Чарлз Спийрман, Торндайк (Бижков, Краевски, 2002). Върху откритията на изследователи като Роберт Ладло и Джон Карол се базира „класическата тестова теория“ и се развива практиката на езиковите тестове в средата на XX век (Шопов, 2013).

В България тестологията е сравнително нова наука. До 40^{-те} години на XX век тестовите са разработвани и прилагани основно за изследователски цели и за подготовка на студенти. Изключителни заслуги за полагането на основите на тестологията има изследователската дейност на Димитър Кацаров и Генчо Пиръв. Значим принос за развитието на тестологията и дидактическите тестове у нас имат учени и изследователи като Енчо Герганов, Георги Бижков, Фелиянка Стоянова, Володар Краевски, Евгения Стоименова, Тодор Шопов, Кирил Банков, Красимир Калинов и редица други. Възникнала като междудисциплинарна наука за създаване на измервателни диагностични методики, тестологията включва в себе си измерване, статистика и математика. Тази наука е множество от математически модели, предлагащи различни решения на проблемите. Тестологията като дял от психометрията се занимава с теорията за конструиране и анализ на тестове. От проучената литература (Bühner, 2011; Бижков, Краевски, 2002; Grotjahn, 2000; Джалев, 2013) може да се направи обобщение, че съществуват две основни тестови теории: класическата тестова

теория (нормативна) и вероятностна (теорията за отговор на тестов въпрос⁵/ пробабилистична⁶). Въпреки някои съществени разлики между тях те имат общи теоретични конструкти като характеристиките трудност и дискриминативна сила на въпросите. Разработената от Кронбах (1963 г.) „генерализираща теория на тестовете“ е опит за преодоляване на недостатъците на класическата теория на тестовете и се приема като преход между класическата тестова теория и вероятностните теории на тестовете, които се развиват и налагат от 60-те години на XX век (Бижков, Краевски, 2002). Появата и развитието на теорията за отговор на тестов въпрос е стъпка напред за оценяването. Тя се базира на въведената от П. Лазарсфелд концепция за латентните черти, чието измерване е индиректно, чрез наблюдение на поведението на индивиди, изпълняващи релевантни задачи (Джалев, 2013).

Според проучванията първите езикови тестове възникват преди Първата световна война в САЩ (Spolski, 1995). Традицията да се разработват езикови тестове е поставена в САЩ с цел тяхното използване за регламентиране на имигрантските потоци. Историческото развитие на теорията на езиковите тестове е разгледано като връзка между психометрията и езиковата теория и е представено в три фази: преднаучна, психометрично-структуралистична и психолингвистично-социолингвистична (Glaboniat, 1998). Според Глабониат социолингвистичните фактори и понятието *комуникативна компетентност* заемат важно място в областта на езиковите тестове едва след „прагматичния обрат“ в обучението. Тази концепция обхваща не само знанията за езиковите правила, а преди всичко прагматичните аспекти на комуникацията, конструкта на компетентността за употреба на езика, който включва също така и знанията за социалната и комуникативната адекватност и усвояването на правилата за уместна употреба на езика в социокултурните рамки (Glaboniat, 1998).

С разрастването на възможностите за професионална мобилност в рамките на Европейския съюз става все по-важно прозрачното измерване и оценяването на езиковите постижения. Европейският съвет от години полага усилия за повече прозрачност и сравнимост при оценяването и сертифицирането на езиковите компетентности. Публикуването на **Общата европейска езикова рамка** допринася за осъзнаването и разбирането на фундаментални въпроси при ученето, преподаването и оценяването на езици. През 1990 година е създадена

⁵ В литературата на английски език се използват термините IRT (Item – Response – Theory) и LTT (Latent – Trait – Theory)

⁶ Терминът *пробабилистична тестова теория* се използва в научната литература на немски език.

ALTE (Association of Language Testers in Europe), която си поставя следните цели: уеднаквяване на езиковите нива за подпомагане на международното признаване на езиковите сертификати в Европа; уеднаквяване на критериите за качество за всички фази при разработването на изпити, т.е. структура на теста, провеждане на изпита, оценяване, сертифициране и съхраняване на данните. ALTE разработва „Наръчник за разработване и провеждане на езикови тестове“, който е допълнение към инструментариума⁷ за подпомагане на всички, които използват ОЕЕР (EUROPARAT/ALTE, 2012). Важно значение в областта на тестването – и оценяването и особено в аспекта на самооценяване на компетентностите и напредъка при изучаване на чужди езици – има **Европейското езиково портфолио** (Europäisches Sprachenportfolio für Jugendliche und Erwachsene, 2001). Усвояването на техниките на работа с ЕЕП помага на учащите да не бъдат зависими само от чуждото оценяване и е подходящият инструмент за документиране на ученето на езици през целия живот и на многоезичието на хората (Glaboniat, Müller, 2006). Поставянето на оценка винаги е по отношение на нещо. Основна цел на ОЕЕР и на другите инструменти на Европейския съвет е да бъдат мярката при оценяването, с помощта на която да могат да бъдат направени валидни, надеждни и сравними изказвания за езиковите знания.

1.2. Видове оценяване и категоризиране на тестовете

От проучената литература може да се направи извод, че съществуват различни видове оценяване и оценки, като всяка от тях има своето предназначение. Класификация на видовете оценяване може да бъде направена, като се отчетат различните фактори: според *начина на интерпретация на резултатите* (нормативно и критериално оценяване), според фактора *кой оценява* (вътрешно и външно оценяване), според фактора *обхват* (индивидуално, групово, училищно, регионално и национално оценяване), според факторите *време на провеждане* и *организация* (текущо и заключително оценяване) и според фактора *начин на оценяване* (аналитично и холистично оценяване).

⁷ Към полезните инструменти могат да бъдат отнесени:

Manual for Relating Language Examination to the European Framework of Reference for Languages: Learning, teaching, assessment (Europarat, 2009); Reference Supplement to the Manual for Relating Examination to the CEFR (Banerjee 2004; Verhelst 2004, Kaftandjieva 2004; Eckes 2009); European Language Portfolio; Content Analysis Grids; Profile Deutsch.

Тестовите могат да бъдат категоризирани според техните качества като *стандартизирани* и *нестандартизирани тестове* (Стоянова, 1996; Glaboniat, 1998; Стойчева, 1998). Стандартизираните (формалните) тестове се разработват и проверяват (апробират) според утвърдените научни стандарти. Приема се, че един тест е стандартизиран, когато неговите качества отговарят на общоприетите критерии за качество (Nissen, 2003). Критериите за качество се отнасят до „качествата на отделните тестови задачи и теста като цяло, допустимата грешка на измерването, вида на използваните измерителни скали и начина на определяне на стандартите за оценка на резултатите от теста.“ (Стоянова, 1996, с.47). Стандартизираните тестове се провеждат според формалните методи – при определени и сравними условия. Характерното за стандартизираните езикови тестове е, че те проверяват отделно в независими тестови части четирите умения: слушане с разбиране, четене с разбиране, писане и говорене (Grotjahn, Kleppin, 2015; Glaboniat, 2010). При стандартизираните тестове се разработват тест-спецификации, които съдържат данни и описания на тест-конструкта, изпитните цели, участниците в изпита, форматите на задачите, вида на инструкциите, броя на задачите, времето за работа и последователността на задачите. Нестандартизираните тестове не отговарят на научните методи и по принцип се използват в учебното ежедневие в училище или университета. Те се създават от учители и техните цели са да предоставят информация за конкретна група и за определено учебно съдържание (Albers, Bolton, 1995). Много често те се използват еднократно, затова усилията и разходите, необходими за създаването на стандартизираните тестове са нецелесъобразни при тяхното разработване. При тях могат частично да се използват разработените в рамките на теорията на теста методи за създаване и оценка на измерителни инструменти (Glaboniat, 1998). Най-важните критерии за качество на нестандартизираните тестове са тяхната валидност по отношение на ситуацията и тяхната съдържателна валидност (Grotjahn, 2000a).

Според вида на тестването и формата на теста могат да бъдат разграничени *изолирано* (discrete-point/ analytic) или *интегрирано/ комбинирано* (integrated) *тестване*. Изолираното тестване си поставя за цел измерването на владенето на специфични изолирани езикови феномени (Grotjahn, 2003). При изолираните тестове акцентът се поставя върху частични умения или върху владенето на определени езикови структури. Комбинираните тестове включват задачи към две или три умения, свързани по различен начин, и си поставят за цел активирането на различни дейности, напр.

дейността *четене* може да бъде свързана с писане или говорене. Комбинираните тестове съответстват много повече от изолираните тестове на „действителното приложение на езика в реални комуникативни ситуации“ и прилагането им в контекста на обучението е от особено значение (Bolton, 1996, с. 103). Глабониат разграничава два метода за измерване на езиковите компетентности: *директното* и *индиректното тестване*. При директния тестов метод поставените задачи отговарят на конкретните изисквания и на измерваните компетентности, като тестовият резултат на изпитваните лица дава възможност за директен извод за техните способности в съответстващи реални ситуации, докато целите и конструктът на индиректното тестване понякога не могат да бъдат разпознати и „резултатите от него често могат да бъдат интерпретирани едва след сложни методи за дешифриране.“ (Glaboniat, 1998, с. 17).

В методическата литература тестове се класифицират принципно като входни тестове⁸, тестове за постижение⁹ и тестове за ниво¹⁰ (Kniffka, 2003). Типично *завходните тестове* е включването на отделни задачи, които имат покачващо се ниво на трудност. Те имат за цел „установяване на входното равнище на комуникативната компетентност на учещите в началото на учебната година или в началото на езиковия курс.“ (Стойчева, 1998, с. 62). *Тестове за постижение* целят контрол за ученици и учители, отчитат какъв е напредъкът и до каква степен са постигнати учебните цели. Тези тестове се отнасят към определени уроци и съответстват като ниво на трудност на учебната прогресия. Резултатите дават информация за напредъка и са ориентир за учителите в учебния процес. *Тестовете за ниво* служат за установяване на езикова компетентност като предпоставка за бъдеща употреба на езика, затова те са ориентирани към речевата употреба. Те проверяват моментното ниво на владене на езика и са критериално ориентирани по предварително определени спецификации и описания на компетентностите (Glaboniat, 1998). Специфично за езиковите тестове е категоризирането им като *тестове заречева употреба* (Performanztests) и *тестове за езикова компетентност* (Kompetenztests), което се базира на разликата между езикова компетентност и езикова употреба (Grotjahn, 2000; Vollmer, 2003).

В доклада „Национални тестове за оценяване на учениците в Европа: цели, организация и използване на резултатите“ различните тестове са отнесени към три групи: *крайни тестове*, *национални тестове* и *тестове за текущо оценяване* (Евридика,

⁸Einstufungstests (placement)

⁹Leistungstests (achievement)

¹⁰Qualifikationstests (proficiency)

2009). Първата група включва крайните (финалните) тестове, които обобщават постиженията на отделните ученици в края на учебната година или в края на конкретен образователен етап. Резултатите от тези тестове се използват за връчване на сертификати или за вземане на важни решения, за избор на училище или за преминаване от един клас към друг. Към втората група се отнасят националните тестове. Тяхната основна цел е извършване на мониторинг и оценка на училищата и/или образователната система. Националните тестове се използват като показатели за качеството на преподаване на учителите и като мярка за ефективността на образователните политики и практики. Третата група включва тестовете за текущо оценяване, които са насочени към подпомагане на учебния процес на отделните ученици. Те имат за цел подобряване на процеса на преподаване и обучение посредством идентифициране на конкретни потребности от учене и приспособяване на преподаването съобразно тези потребности (Евридика, 2009). В обобщение може да се каже, че разграничаването на видовете оценяване (изпити) и видовете тестове се прави в зависимост от целта, за която ще бъде използвана събраната от тях информация за нивото на знания на изпитваните лица (Perlmann-Balme, 2001).

1.3. Функции на оценяването на постиженията в езиковото обучение

Темата за оценяването често предизвиква негативни чувства. Тестовете и изпитите натоварват климата на преподаването и ученето, защото в много училища те задават такта на учебните занятия. Често селективната функция се покрива или частично съвпада с развиващата функция, което става причина тестовете да имат нежелан страничен ефект, а именно – стесняване на учебните цели, като често се учи само това, което е в изпитните материали (Perlmann-Balme, 2001). Оценяването чрез тестове и изпити изпълнява различни функции в чуждоезиковото обучение. Изследователите на обучението по чужд език разглеждат функциите на изпитването, тестването и оценяването от три различни перспективи: от перспективата на преподавателите, от тази на учещите и от образователно-политическа перспектива (Grotjahn, Kleppin, 2015).

От перспективата на преподавателите функциите на изпитването, тестването и оценяването са: **диагностициране** на специфични силни и слаби страни на учениците; **подпомагане на учениците**; **определяне на входно ниво** – което дава информация за това на кое езиково ниво съответстват постиженията на учениците; **идентифициране**

на прогрес, при което учителят получава информация колко е напреднал всеки един ученик или групата като цяло; **установяване на постигнатите учебни цели** в съответствие с образователните стандарти и нормите в учебните програми; **селекция** с цел следващо образование или работа; **класиране** на учениците в рамките на класа/групата; **поставяне на текущи оценки** на учениците; **сравняване на постиженията** на ученици от различни класове/групи; **обратна връзка** за учители, ученици и други заинтересувани лица, като например родители (в резултат на оценяването учениците получават описание на компетентностите, които са усвоили, а учителите могат да използват резултатите от оценяването като повод да преосмислят учебните цели, а в някои случаи и да направят промени.); **мотивиране на учениците** (при оценяване учениците са подтиквани да положат повече усилия, да затвърдят определени знания или да преговорят конкретно съдържание.); **дисциплиниране** (понякога оценяванията се използват с цел учениците да бъдат доведени чрез една слаба оценка до това да се държат адекватно в учебните часове.); **развиване на компетентностите за самооценка** като важна предпоставка за изграждането на автономност у учениците.

Учениците също имат различни причини да се интересуват от оценяването. От тяхна перспектива изпитването, тестването и оценяването могат да имат следните функции: **информирание за индивидуалното ниво на постижения** чрез тестове за самооценка; **установяване на личния прогрес** чрез самооценка в различни моменти от учебния процес; **развиване на лични компетентности за диагностика и рефлексия**; **получаване на важен сертификат**, удостоверяващ нивото на владее на езика.

Ако изпитването, тестването и оценяването се разгледат от перспективата на образователната политика, могат да бъдат формулирани други функции: **сравняване на резултатите** на училища и региони от проведените национални или регионални външни оценявания; **измерване на постигането на целите** по отношение на основните компетентности в съответствие с ДОИ за учебно съдържание и учебните програми; **положително влияние върху обучението**.

Оценяването може да има значими последствия за учениците, което е достатъчен аргумент за осъзнаване на важността на точното определяне на целите и функциите на един изпит и какви ще бъдат влиянията от него. Създаването на тестове предполага познаването на критериите за оценка на качеството на тестовете и тестовите задачи.

1.4. Критерии за качество на езиковите тестове и тестовите задачи

Тестовите играят важна роля при оценяването на постиженията и от това произтича необходимостта от поставянето на високи изисквания към качеството на измерването. В проучената литература като основни критерии за оценка на качеството на тестовите и тестовите задачи са посочени *обективност, надеждност и валидност* (съдържателна, критериална, конструктна). Към тях се добавят критерии *сравнимост, балансираност, специфичност, икономичност, практичност, дискриминативна сила на задачите, стандартизация, обратно влияние върху обучението, автентичност, полезност* и редица други (Трим и др., 2006; Стойчева, 1999; Бижков, Краевски, 2002; Стефанова, 2015; Grotjahn, 2000, 2006, 2008; Bachman, Palmer, 1996; Bachman, 1996; Glaboniat, 1998; Ingenkamp, 1997; Lienert, Raatz 1998; Perlmann-Balme, 2001, 2010; Sanders 2000; Bühner, 2011; Rost, 2004; Grotjahn, Kleppin, 2015). Разгледани са различни аспекти на валидността като *контекстна валидност* и *когнитивната валидност*, както и специфичните критерии за качество на тестове за оценяване на езикови компетентности: *автентичност, интерактивност, прозрачност* и *ефект на обратно влияние* върху съдържанието на учебния процес и обществените развития. Представен е разработеният от Дласка и Крекелер каталог с критерии, приложими при създаване на тестове, предназначени за употреба в учебните занятия (Dlaska, Krekeler, 2009).

Наред с критериите за качество на тестовите и изпитите, е отделено внимание и на типичните грешки при тестването и оценяването. Представени са често допусканите грешки при разработване на тестове по отношение на *прозрачността* (неразбираемост на указанията към задачите, необявяване на критериите за оценяване); на *валидността* (не се проверява това, което трябва да се проверява); на *вариативността на методите* (използват се едни и същи тестови формати); на *диференциацията, практичността* и *релевантността* (Mendez, 2013). В изложението са включени методи за справяне с грешките при разработване на задачи: критична саморефлексия, колегиална обратна връзка, съвместно разработване и обсъждане на задачи и взаимно коригиране (Mendez, Münchbach-Wellner, 2013). Разгледан е и *проблемът за едностранността*, който не може да бъде преодолян напълно, но може да се редуцира чрез осъзнаването му от изпитващите лица (Bolton, Perlmann-Balme, 2006).

1. 5. Етапи при създаването на тестовете и видове тестови формати

В дисертацията са представени етапите, през които преминават дейностите по създаването на тестове за стандартизирани оценявания: определяне и конкретизиране на целта на измерването; съставяне на тест-спецификация; създаване на тестови задачи; апробация на теста; анализ на качествата на отделните задачи; втора апробация; анализ на качествата на теста след втората апробация; определяне на стандартите за оценка; подготовка на теста за масова употреба. Спазването на основните етапи при разработването на тестовете се приема като „гаранция за постигане на стандартите за качество.“ (Стоянова, 1996, с.58).

Дефинирано е понятието *тестов формат*¹¹, което се отнася до начина, по който трябва да се реагира на указанията, и как трябва да се отговаря на поставените въпроси (Glaboniat, 1998). Направено е уточнението, че различните видове тестови формати поставят различни изисквания и форматите на отговор са с различна степен на трудност. На ниво тест-формат в центъра стои „документирането на разбирането“, начинът, по който разбраното да бъде демонстрирано и да стане измеримо. (Glaboniat, 1998, с. 93) От проучената литература се прави заключение за основното разграничаване на тестовите формати между задачи, чието вярно решение се избира от зададени варианти на отговор (формати с избираем отговор, затворени), и задачи, чието решение се конструира самостоятелно (формати със свободен отговор; отворени; конструктивни формати) (Стоянова, 1996; Glaboniat, 1998; Lienert, Raatz, 1998; Rost, 2004). За целите на конструктивната област се приема и разграничението между рецептивни или затворени формати; репродуктивни или полуотворени формати и продуктивни или отворени формати (Glaboniat, 1998).

В тази част на изследването е представена типологията на *форматите с избираем отговор (рецептивните формати), полуотворените формати на отговор (репродуктивните формати) и форматите със свободен отговор (продуктивните формати)* и са систематизирани техните предимства и недостатъци по отношение на проверката на рецептивните и на продуктивните умения. Разгледана е уместността на тестовите формати в зависимост от проверяваната компетентност, езиковото ниво на целевата група и целта на проверката, като са направени изводи, че няма добър или лош формат. Изборът на представените теми произтича от убеждението, че знанията за теорията на теста са важен фактор за развиване на уменията в сферата на оценяването.

¹¹ В научните публикации на български език най-често употребяван е терминът „вид на задачата“.

Глава 2.

Методи за разработване и оценка на тестови задачи за оценяване на езикови компетентности

2.1. Принципи за създаване на комуникативни езикови тестове

В тази част от изследването са представени основните принципи и методи за създаване на комуникативни тестове. В тяхната основа стоят моделите на комуникативната компетентност. Разработеният от Бахман и Палмър модел на комуникативна компетентност повлиява емпиричните изследвания и редица проекти за разработване на тестове се опират на него при дефинирането на конструкта *комуникативна езикова способност* (Bachman, Palmer, 1996). Техният модел се основава на предложението от Канал и Суейн (1980) модел, който включва като първи елемент *комуникативната компетентност*, към която се отнасят лингвистичната, социолингвистичната и стратегическата компетентност и като втори елемент – *комуникативното поведение* (Димитрова-Гюзелева, 1999). При дефинирането на комуникативната езикова компетентност Бахман и Палмър разглеждат компонентите *езикова компетентност, стратегическа компетентност и психофизиологична компетентност*. Езиковата компетентност те определят като специфична, съхранена в паметта езикова наличност на знания, а стратегическата компетентност описват като генерална метакогнитивна способност за прилагане на езикова и неезикова наличност на знания и способности в дадена ситуация. Стратегическата компетентност е от основно значение по отношение на употребата на езика по време на тестова ситуация и извън нея. От психофизиологична гледна точка езиковият процес може да се раздели на две полета. Първото поле се отнася до комуникационния канал: *визуално* (писмено) и *слухово* (устно), а второто обхваща начина на използване на езика, възможностите за *рецепция* и *продукция* (Bachman, Palmer, 1996). Към тези аспекти в модела на ОЕЕР са добавени понятията *интеракция* и *трансакция* (Трим и др., 2006). В ОЕЕР комуникативната компетентност е разделена на три компонента: *лингвистичен, социолингвистичен и прагматичен компонент*. По принцип става въпрос за същите аспекти, които разглеждат Бахман и Палмър, но са използвани различни форми на описание. Комуникативната компетентност се изразява в рамките на посочените полета, при което към всеки начин на използване могат да бъдат подредени различни прагматични езикови дейности. Трудности има при подреждането на езиковите

дейности, които обхващат няколко области, напр.: продуктивните дейности, които включват и интеракция – като разговорите при устното изразяване. Подобен проблем се съдържа и в дефиницията на отделните езикови дейности (области на умения). От една страна е стремежът да се отчете сложното взаимодействие на различните дейности и частичните умения (лексикална, граматическа, фонетична, правописна компетентност), а от друга страна – намирането на практична система за описание. Общата цел на тези два модела е преодоляването на традиционното делене на четири области на комуникативни умения, което не е целесъобразно и автентично, ако се разгледа в аспекта на тяхната изява в реалността. Въпреки това в практиката на изпитването се е запазило формалното разделяне на тестовете на субтестове според четирите умения.

Първото изискване при разработването на комуникативните тестове е свързано с определянето на проверявания конструкт. Следвайки описаните модели на комуникативна езикова компетентност, този конструкт трябва да обхваща *езиковите компетентности* (организационна и прагматическа), които съдържат граматическа, текстова, функционална, социолингвистична компетентност, както и *стратегическата компетентност* (компенсаторни стратегии, дискурсни стратегии). Комуникативните тестове трябва да включват имплицитно или експлицитно проверка на всички частични компетентности (Glaboniat, 1998).

Изводът от проучената литература е, че изказвания за езиковите умения могат да бъдат направени в реални ситуации, когато изпитваното лице докаже уменията си в съответстващи ситуации. Следователно тестовите ситуации трябва да са възможно най-близки до реалната ситуация на използване на проверяваното умение. *Принципът за автентичност* при проверка на рецептивните умения се отнася не само към компонента *текст* (използване на автентични текстове), но и към целта на рецепция и към очакваната реакция – следователно към тестовия формат (Glaboniat, 1998). Другият важен принцип е *ефектът на обратна връзка*. Той се определя от въздействието, което тестовете и изпитите оказват върху учебните занятия и учебните материали, и може да бъде както положителен, така и отрицателен. Пряка връзка със съдържанието на тестовете има типологията на задачите и тестовите формати, които играят важна роля. В тази връзка един от принципите на комуникативните тестове е включването на разнообразни *директни тестови формати*. (Edelhoff, 1985; Glaboniat, 1998). В проучената литература има много примери за влиянието на афективните фактори върху резултатите от тестовете – сред тях са изпитната атмосфера, страхът, нервността,

мотивацията. От изискването изпитваното лице да бъде поставено в центъра на интереса се извежда принципът *насоченост към учащите*. Въздействие върху изпитния резултат могат да окажат различните нагласи и предварителните очаквания към тестовете. Ролята на *интеракцията* има важно значение както при устното изразяване, така и при писменото изразяване. Комуникативните тестове вземат под внимание компонентите на интеракция при определяне на комуникативните цели в тест-спецификацията и при оценяването на езиковите постижения (Glaboniat, 1998).

2.2. Оценяване на уменията четене с разбиране

В методиката на обучението по съвременни езици се приема, че „делта на обучението по четене е обучаваният да се научи да извлича информация, която се съдържа в даден текст на изучавания език.“ (Стефанова, 2015). В езиковото обучение четенето с разбиране се определя като комплексна дейност. В сравнение с писането, където има конкретен продукт и на базата на текст се правят изводи и се оценяват знанията и уменията на тествания, при четенето са невъзможни директните наблюдения. Най-често действията и стратегиите на четящия остават скрити за външния наблюдател (Alderson, 2000). Проверяващият не може да проследи протеклите при четенето когнитивни процеси – те зависят от индивидуалните особености на отделното лице. Следователно при създаването на тест за *четене с разбиране* се спазва принципът: задачите/въпросите към текстовете да активират конкретни когнитивни и текстообработващи операции, които са предмет на оценяване. Според Арас тестът не представя реални езикови действия, а ги отразява. Следователно на базата на резултатите могат да бъдат направени предположения за действителната компетентност на тествания или по-точно за това как тестваното лице действително би реагирало в една реална ситуация. Затова е необходимо признаците на реалната езикова ситуация да се възпроизведат възможно най-точно в теста чрез иницирираните езикови действия, като проверяваните компетентности се обхвалят под формата на стратегии за четене и когнитивни процеси (Arras, 2006).

Разработването на тестове за *четене с разбиране* се базира върху следните аспекти: стратегиите за четене и разбиране, които ще бъдат измервани; съдържателното (тематичното) и езиковото оформяне на текста; вида на текста; формата на задачата, с помощта на която се оценява разбирането, и метода за оценяване. Определянето и

контролирането на степента на трудност на един тест зависи от различни фактори. Гротян предлага следните детерминанти за определянето на степента на трудност:

- фактори, които определят степента на трудност на текста;
- фактори, които определят степента на трудност на задачите;
- фактори, които определят връзката между текст и задачи (Grotjahn, 2000b).

Чрез комбинираното отчитане на тези фактори е възможно по-точно определяне степента на трудност на един тест за *четене с разбиране*. Трудността на задачите за *четене с разбиране* се определя и от трудността на инструкцията за изпълнение на задачите, както и от скалата за оценяване. Трудната задача става по-лесна, когато оценяването е по-благоприятно – и обратното (Projektgruppe TestDaF, 2000).

За решаването на една комуникативна задача, което означава – за постигането на конкретна комуникативна цел, учащият използва както своите комуникативни езикови компетентности (лингвистични, социоллингвистични, прагматични), така и компетентности като знания за света и социокултурни знания. От една страна, има различни възможности и стратегии да бъде постигната една и съща комуникативна цел, но от друга страна, важна роля играят различните индивидуални фактори като предварителните знания. От проучената литература са обобщени критериите, определящи трудността на комуникативните задачи, представени на фигура № 1.

Фигура № 1 Критерии за определяне нивото на трудност на комуникативните задачи



При определянето на трудността на задачите трябва да се вземат предвид различните фактори, които са включени в процеса на комуникация. За практическата работа по създаването на тестове за оценяване на уменията *четене с разбиране* е важно разглеждането на следните въпроси: къде могат да се намерят добри изходни текстове; на какви критерии трябва да отговарят тези текстове; какви задачи са подходящи за тези текстове (Bolton, Perlmann-Balme, 2006).

2.3. Връзка между вида на текста, вида на задачата и целта на четене

Понятието *текст* „се използва за обозначаване на всички езикови продукти, които ползвателите на езика възприемат, произвеждат или обменят, независимо дали е устно изказване, или нещо написано.“ (Трим и др., 2006). Видът на текста предполага прилагането на различни стратегии (стилове) на четене. При работа с текстове от съществено значение е каква цел на четене се следва, защото целта определя избора на стратегия (стил) на разбиране (Bolton, 1996; Neuneru.a., 2009). Целите на четене и стиловете на четене могат да бъдат:

- да се разбере и извади основната, най-важната информация – глобално разбиране или курсорно четене;
- да се потърси, разбере и извади определена информация – селективно разбиране;
- да се разбере всеки детайл (подробностите) в текста – детайлно разбиране.

Най-често целта на четене зависи от прагматичните функции на съответния вид текст (Neuneru.a., 2009). Тя определя поведението на читателя, респективно стила и стратегията на четене. В зависимост от вида на текста се прилагат различни стратегии на четене. Различните стратегии се проверяват чрез различни видове тестови задачи. Не всеки вид текст е подходящ за всеки вид задачи. Следователно при съставянето на тестове за проверка на уменията за разбиране при четене¹² трябва да се уточнят на първо място целта на четене и видът на текста. Видът на текста предполага използването на съответстващия стил и стратегия, а стратегията обуславя вида на тестовите задачи. Тази взаимовръзка е разгледана в редица изследвания, посветени на оценяването на уменията за *четене с разбиране* (Стойчева, 1999а; Bolton, 1996). Въз основа на проучената литература тези взаимовръзки са обобщени и систематизирани в настоящата работа с цел да бъдат взети под внимание при разработването на тестове за проверка на уменията *четене с разбиране*.

¹² Условието е валидно и при съставянето на тестове за проверка на уменията *слушане с разбиране*.

2.4. Текстовете за тестови задачи за оценяване на уменията *четене с разбиране*

Учебната цел – развиване на комуникативните умения на чуждия език в ежедневни ситуации, може да бъде постигната и съответно проверена чрез симулация на близки до реалността ситуации и чрез използването на автентични материали за четене на целевия език. Приемайки, че изборът на подходящи текстове е важна задача при съставянето на тестове за проверка на езиковите компетентности в различните области на умения, в тази част на дисертацията са систематизирани критериите за подбор на текстове за тестови задачи за проверка на уменията *четене с разбиране*. Въз основа на проучената литература са обобщени следните важни изисквания: текстовете да са автентични или дидактизирани, съдържащи типичните за текста характеристики; обемът на непозната лексика да не е голям; ключовите думи да са позната лексика; синтактичната структура да отговаря на съответното езиково ниво; обемът и структурата на текстовете да съответстват на предвиденото в учебното съдържание; текстовете да съответстват на жизнения опит, степен на развитие и интереси на целевата група; текстовете да са интересни и актуални, съдържащи разнообразни информации (Стойчева, 1999а; Glaboniat, 1998). Разгледан е факторът *трудност на текстовете* за оценяване на уменията *четене с разбиране*. Нивото на трудност на текстовете се определя въз основа на анализа на различни критерии (Glaboniat, 1998). Посочените от Глабониат критерии са систематизирани и представени на фигура № 2.

Фигура № 2 Критерии за определяне на нивото на трудност на текст



Най-често използваните източници за автентични текстове са статии от вестници и списания, брошури, статии от интернет и рекламни материали. За целите на практическата работа по разработване на тестове е представен методът **съставяне на карта на текста**¹³ (Textmapping), който е една от процедурите, предхождащи работата по съставянето на задачи/въпроси към текстове. Целта на метода е преди създаването на задачите да се конкретизират най-ясните, най-често възприеманите съдържания на един текст, към които впоследствие да бъдат създадени задачи. Методът повишава валидността на задачите, като се базира на предварителното определяне на целите и стила на четене въз основа на характера/вида на текста.

2.5. Правила за създаване на тестови задачи с избираем отговор

Според ОЕЕР тестовите задачи не трябва да се различават от изпълняваните в ежедневието задачи и по своята същност те могат да бъдат: когнитивни, ориентирани към уменията, творчески, рутинни, несложни или комплексни, еднодименционални или обхващащи поредица от подусловия. Форматите с избираем отговор са най-често използваните при проверка на уменията *слушане с разбиране* и *четене с разбиране*, а добрите задачи с избираем отговор се създават трудно. Въз основа на проучената литература (Стоянова, 1996; Бижков, 1996; Стоименова, 2000; Банков, 2012; Glaboniat, 1998; Grotjahn, Kleppin, 2015) са систематизирани изискванията, които трябва да бъдат спазвани при създаването на задачи с избираем отговор. Представените правила се отнасят към съставянето на задачи от рецептивните и репродуктивните формати за проверка на уменията *четене с разбиране*.

Инструкциите са важна съставна част от тестовите задачи. При формулирането на **инструкциите за изпълнение на задачите** трябва да се имат предвид аспектите при процеса на четене, нивата на преработка на информацията и стилът на четене. Инструкциите трябва да са формулирани прецизно и разбираемо (да се използват лексика и синтаксис, подходящи за езиковото ниво); да са ясни и кратки и да съдържат следните информации: какво трябва да се извърши, в каква последователност и за какво време.

Общите **правила за съставяне на тестови задачи с избираем отговор** се отнасят до различни аспекти. Формулировката на условието на задачата трябва да е ясна,

¹³ Виж интернет източници: Textmapping als Methode zur Erstellung von Testaufgaben. http://www.kmk-format.de/material/Fremdsprachen/4-4-7_Textmapping_fuer_Erstellung_von_Leseaufgaben.pdf.

кратка, точна и да не е сложна. При формулирането на условието да не се използват директно фрази от текстовете. При възможност формулировка на условието да е въпросително изречение. Важно е в условието на задачата да са включени основната идея, проблемът, върху който е фокусирана задачата. Задачите към един текст трябва да са независими една от друга: решението на една задача да не дава отговор на друга задача. Да не се допуска формулирането на задачи, които могат да бъдат решени само с помощта на общи знания, без да се чете текстът. Да се избягват негативните фрази в условията на задачите, а когато се налага тяхното използване, отрицанието да бъде маркирано. Задачите да имат само един еднозначно верен отговор и два (или повече) еднозначно грешни отговори. Визуалното оформление на задачата да съдейства за нейното по-лесно възприемане. При създаването на повече задачи да се внимава на какви позиции са верните отговори с цел да не са на една и съща позиция или случайно да е възникнала повторемост. Важно е също така решението къде да бъдат разположени задачите – преди или след текста за четене. Разполагането на задачите може да окаже влияние върху тяхното обработване. Позиционирането им трябва да бъде съобразено с тест-конструкта (стила на четене). Например селективният стил на четене предполага задачите да са разположени преди текста.

Систематизираните **правила за създаване на варианти за отговори към задачите с избираем отговор** са насочени на първо място към общите изисквания. Те включват следното: **вариантите за отговор** трябва да са ясно и еднозначно формулирани; да са с приблизително еднаква дължина; да са хомогенни по своя характер; да са естествено допълнение на изходното изречение в условието на задачата; да не съдържат повтарящи се думи и изрази; да не съдържат ключови думи или фрази от текста, подсказващи верния отговор; да не съдържат грешки (езикови, фактологически); да са формулирани по-достъпно от формулировките в текста; да са подредени логично по последователността на своето появяване в текста и хронологично – когато съдържат числа. Представени са и конкретни **изисквания към верните отговори**. До верния отговор трябва да се стига на базата на знания и разбиране на текста, а не въз основа на логиката. Верният отговор трябва да се съдържа в текста, а неговата позиция да е случайно подбрана. Систематизираните **изисквания към дистракторите** са определени от техните важни функции. Дистракторите трябва да са смислени алтернативни отговори, за да изпълняват функцията си да „разсейват“. Те трябва да бъдат логични и правдоподобни, като тяхното съдържание се базира на информация от

текста. Дистракторите трябва да са възможни отговори за тези, които не притежават знания и умения да решат задачата, и да не са възможни отговори за тези, които притежават знания и умения да решат задачата. Дистракторите могат да бъдат типични грешки, които се допускат от учащите.

Спазването на изискванията за създаване на тестови задачи е важно условие, защото качеството на всеки тест зависи от качеството на задачите в него. Качествените задачи са подчинени на конкретна комуникативна цел и са насочени към потребностите на учащите извън учебната ситуация, като симулират близки до реалния живот изисквания.

Представените критерии за експертна оценка на текстове и на тестови задачи с избираем отговор и с полуотворен формат на отговор за оценяване на уменията *четене с разбиране* изхождат от правилата за разработване на комуникативни тестове, които са утвърдени във водещите европейски тестови центрове. Те са изведени въз основа на анализа на проучените източници (Стоянова, 1996; Бижков, 1996; Стоименова, 2000; Банков, 2012; Glaboniat, 1998; Grotjahn, Kleppin, 2015; EUROPARAT/ALTE, 2012).

В дисертационния труд са представени следните критерии за експертна оценка:

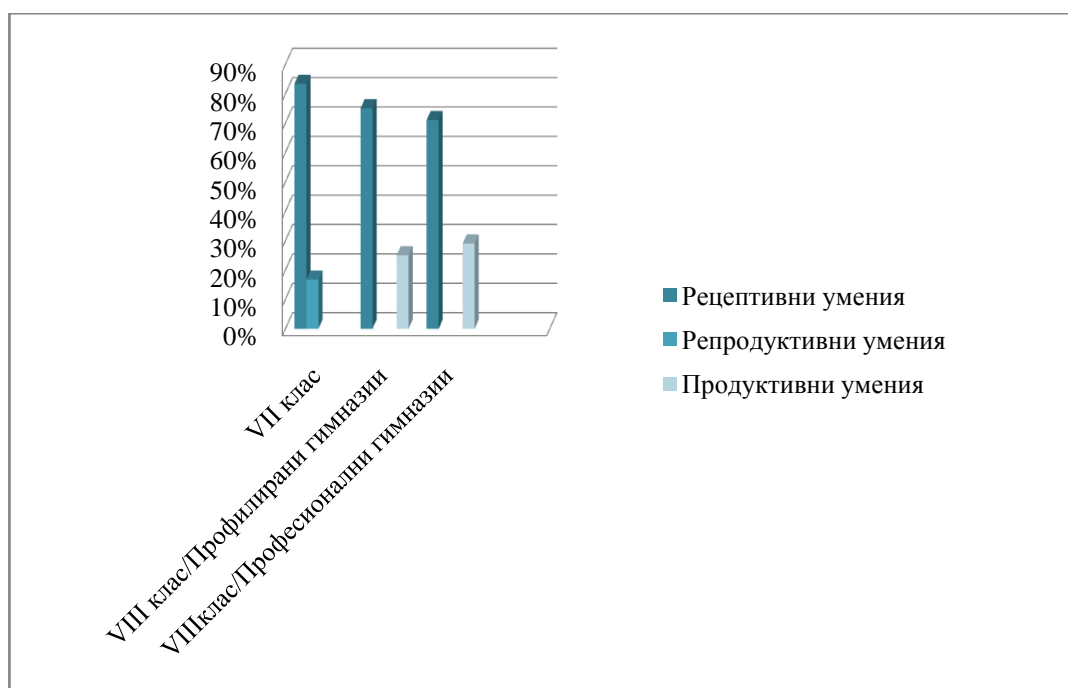
- Критерии за експертна оценка на текст и задачи с избираем отговор към компонент *четене с разбиране*:
 - критерии за оценка на характеристиките на текста за четене;
 - критерии за оценка на инструкцията за изпълнение на задачите към текста;
 - критерии за експертна оценка на задачи с избираем отговор.
- Критерии за експертна оценка на текст и задачи с кратък отговор (полуотворен формат на отговор) към компонент *четене с разбиране*:
 - критериите за оценка на характеристиките на текста за четене;
 - критерии за оценка на инструкцията за изпълнение на задачите към текста;
 - критерии за оценка на характеристики на задачите с кратък отговор (полуотворен формат на отговор) към текст за четене.

Главната цел на предложените критерии за експертна оценка е подпомагане на работата по създаване на тестове на преподаватели, експерти и всички, свързани с оценяването на езикови компетентности.

2.6. Анализ на тестове

В тази част от изследването е направен анализ на структурата на изпитните материали по чужди езици за външно оценяване в VII клас и VIII клас и проверка на тяхното съответствие с наложените критерии за комуникативно ориентирани изпити от водещи европейски езикови организации, членове на ALTE. Тестовете са разгледани в следните аспекти: съотношение между измерването на рецептивни умения и на продуктивни умения; вид на задачите; вид и обем на текста в задачата за писмено изразяване. На първо място е анализирано съотношението на задачите, измерващи рецептивните и продуктивните умения, представено на фигура № 3.

Фигура № 3 *Рецептивни и продуктивни умения*



След анализ на процентното съотношение на задачите, измерващи рецептивните и продуктивните умения, може да се направи следното обобщение: съотношението между измерването на рецептивните и на продуктивните умения е в полза на проверка на рецептивните умения на учениците. От компонентите на комуникативната компетентност на учениците, дефинирани в ДООИ, в настоящия формат на тестовете по чужди езици за външно оценяване се дава превес на задачите за проверка и оценка на рецептивните умения на учениците да извличат информация от текстове при слушане и при четене, както и на задачите за оценка на лексикалните и граматическите знания. В това отношение настоящият формат на тестовете не може да даде ефективна

информация за нивото на постиженията на учениците по отношение на заложените цели в учебните програми по чужди езици за VII клас и за VIII клас, където очакваните резултати са конкретизирани в четири ядра: четене, слушане, говорене, писане.

Относно вида на задачите, наложил се от години формат на тестовете за външно оценяване по чужди езици предлага задачи с избираем отговор (с множествен избор), което е в противоречие с критерия *балансираност*. Оценяването чрез задачи с избираем отговор без съмнение се отразява върху преподаването и в процеса на обучение се акцентира върху запаметяването на езикови структури и върху рецептивните дейности.

Друг важен аспект в анализа е обемът на текста, който учениците трябва да създадат в задачата за писмено изразяване. Видът и обемът на текстовете, които изучаващите чужд език трябва да могат да създават, трябва да са съобразени с критериите за писмено изразяване към съответното езиково ниво, определени в ОЕЕР. В настоящия формат на тестовете за VIII клас в част „Писмено изразяване“ е определено да се създаде един писмен текст с обем 170 – 190 думи (за профилирани гимназии) и 120 – 140 думи (за професионални гимназии). Това налага ограничението да бъдат формулирани задачи, изискващи създаването на няколко вида текстове. В тази връзка е необходимо да бъдат направени и обсъдени нови предложения за задачите за писмено изразяване, чрез които да бъдат оценени уменията за комуникация при писмено общуване, съобразени с нормите на езиковото ниво и отнасящи се към реалистични житейски ситуации.

Анализът на тестовете показва, че те са концентрирани върху онези области на езиковото поведение, в които определено крайно поведение на ученика може да бъде операционализирано и оценено чрез решаване на задачи от форматите с избираем отговор. Изследванията показват, че изпитите могат да променят учебните занятия. Ако тестовите формати съдържат разнообразни задачи за писмено и устно изразяване, то продуктивните дейности ще станат съществена част от учебния процес.

Разгледан е също така проблемът за мотивацията за учене и нагласите към оценяването на постиженията. От проучената литература са направени изводи, че нагласите към оценяването са противоречиви: оценяването може да бъде мотивиращ фактор, но често то действа демотивиращо (Solmecke, 1983; Storch, 2009; Arbinger, Herbert, Eberhard, 1977; Düwell, 1979; Perlmann-Balme, 2006; Butzkamm, Plum, 2006; Bimmel, 2006; Попандонова, 2009). Информацията за влиянието на оценяването върху мотивацията за учене е от съществено значение и важно условие за успешното учене.

Глава 3.

Експериментална проверка на качествата на тестови формати с избираем отговор

Експерименталната проверка включва апробация на тест по немски език с цел изследване на качествата на различни тестови формати с избираем отговор за проверка на уменията *четене с разбиране* по отношение на тяхната надеждност, трудност и дискриминативна сила. Целевата група на експеримента са ученици, които учат немски език в VIII клас в профилирани гимназии и профилирани паралелки на СОУ с интензивно изучаване на чужд език. В разработения тест са включени следните видове тестови формати: задачи с алтернативен отговор, с множествен избор и за съответствие. Въз основа на анализа на данните са направени изводи за тяхната приложимост за постигане на практическите цели на оценяването на уменията за четене с разбиране. Целта на анкетното проучване, проведено в рамките на експеримента със същата целева група, е да предостави информация за нагласите на учениците към оценяването и отношението им към различните видове тестови задачи.

3.1. Представяне на процеса на разработване на теста за експеримента

Разработването на теста за експеримента премина през следните етапи:

- създаване на тест-спецификация;
- избор на текстовете за четене;
- съставяне на карти на текстовете за глобално, селективно и детайлно разбиране;
- редактиране на текстовете и разработване на задачи към тях;
- формулиране на указанията за изпълнение за задачите;
- експертни оценки и преработване на задачите в съответствие с препоръките;
- втора експертна оценка, одобрение на тестовите задачи и подготовката за апробацията на задачите.

При разработването на теста са спазени критериите за подбор на текстове за тестови задачи, процедурата по съставяне на карти на текстовете (textmapping), правилата за създаване на тестови задачи с избираем отговор и за формулиране на указания за изпълнение за задачите, критериите за експертна оценка на текстове и тестови задачи, описани във втора глава.

3.2. Представяне на проведения експеримент: изготвяне на извадка, организиране и провеждане на апробацията на теста

Извадката е изготвена по начин, който отразява адекватно структурата на целевата група – ученици, които учат немски език в VIII клас в профилирани гимназии и профилирани паралелки на СОУ с интензивно изучаване на чужд език. В апробацията взеха участие 209 ученици от 9 паралелки от 6 профилирани гимназии и профилирани паралелки на СОУ с интензивно изучаване на немски език. По средния успех на учениците се оформиха 3 групи участници: ученици с успех *над средното*, ученици със *среден успех*, ученици с успех *под средното*. Всяка от трите групи е представена със среден брой по 69 / 71 / 69 ученици, с което се приема, че извадката е изпълнена.

3.3. Статистическа обработка и анализ на резултатите от теста

Анализът и описанието на резултатите съдържат два важни компонента – статистическа обработка на данните и анализ на получените резултати (статистиките на тестовите въпроси). За целите на изследването са използвани класическата тестова теория като система за оценяване на латентни черти при анализа на тестовите данни, статистическите пакети за обработване на данни SPSS¹⁴ и TiaPlus¹⁵, фокусирани към класически анализ на тест и тестови задачи. Чрез методите на класическата теория на теста са проверени и оценени следните характеристики на тестовите въпроси и на теста като цяло: индексите на трудност и на дискриминативна сила, както и характеристиките на отговорите. При обработката на резултатите от апробацията на теста задачите са изследвани самостоятелно и в три групи, условно наречени формати. Задачите са групирани в три формата според техния формат на отговор. Формат 1. включва задачите с алтернативен отговор (1 – 6) от първата част на теста. Към формат 2. се отнасят задачите с множествен избор (7 – 13) от втора и трета част на теста. Във формат 3. са включени задачите за съотнасяне (14 – 20) от четвъртата част на теста.

От основните психометрични характеристики на теста и на субтестовите и от данните за честотно разпределение на точките в теста могат да бъдат направени следните изводи:

¹⁴SPSSStatistics. SPSSInc., 1997.

¹⁵TiaPlus, UserManual, CITO, 2011, M. @ R. Department, Cito, Arnhem, NL.

- Резултатите показват отлична надеждност на теста като цяло за практически цели: коефициент GLB е 0,84. Тази стойност позволява да се правят изводи за цялата извадка и да се сравняват резултатите по групи (пол) и между отделни ученици.
- Коефициентите GLB за отделните части на теста (за формат 1 – 0,69; за формат 2 – 0,58; за формат 3 – 0,78) са достатъчно високи, за да бъде разглеждан всеки от тях самостоятелно.
- Трудността на теста като цяло е 61,24%. Тази стойност показва, че възпроизвеждането/прилагането на знанията и уменията, измервани с включените в него тестови задачи, е на границата между среднотрудни и лесни за учениците.
- Средният тестов бал е 12,25 точки, което е показател, че учениците, покрили този бал, се справят успешно с 60% от проверяваното учебно съдържание.
- Медианата е 13 точки: над 50% от учениците имат 13 и повече точки. Те са покрили 65% и повече от проверяваното учебно съдържание.
- Модата (балът с най-висока честота) е 13 точки. Учениците с този резултат са покрили 65% от проверяваното учебно съдържание.
- Модата и медианата съвпадат и са близки до средния тестов бал, което е доказателство за симетрично разпределение (Стоименова, 2000). А това означава, че теоретичното разпределение е близко до нормалното.

В анализа на качествата на теста са представени основните психометрични характеристики на отделните задачи от теста, определящи тяхната ефективност: трудност, дискриминативна сила на задачите и дискриминативна сила на дистракторите. Коефициентът на трудност на отделните задачи от апробирания тест варира от 33% до 84%, което е в съответствие с изискванията за ниво на трудност при критериалните тестове: стойностите на **R** да са в интервала [5% – 95%] (Стоянова, 1996). От психометричната обработка на резултатите от апробираните задачи може да се направи извод, че задача № 3 е с много ниска дискриминативна сила (- 0,01). Чрез методите на статистическото изследване не може да определи причината за ниската диференцираща стойност на тази задача. В такива случаи е препоръчително да се направи експертен анализ и оценка на задачата. След експертен анализ на задачата се достигна до извод, че формулировката на задачата не е коректна и отговорът не може да бъде категоричен. Следвайки правилата за селекция на задачи, тази задача е

изключена от теста и от анализа поради ниско качество. След обработката на данните от 19 задачи резултатите показват, че тестът съдържа: 8 задачи с допустима дискриминативна сила (препоръчан долен праг $R=0,20$); 8 задачи с добра за практически цели дискриминативна сила (от $R=0,31$ до $R=0,40$) и 3 задачи с много добра и отлична за практически цели дискриминативна сила (над $R=0,41$). При оценката на дискриминативната сила на задачите се съобразяваме с предложените стойности за дискриминация от К. Банков и Е. Стоименова (Банков, 2012; Стоименова, 2000).

От данните за дискриминативната сила на дистракторите може да се направи извод, че при задачите от формат 1. и формат 2. недостатъчно ефективен е дистрактор А) на задача 8., но тъй като тази задача е с добри показатели по отношение на трудност и дискриминативна сила на верния отговор, тя може да бъде определена като добра за практически цели. При разглеждането на дискриминативната сила на дистракторите при *задачите за съотнасяне* от формат 3. (част 4. от теста) трябва да се вземе предвид същността на този тестови формат. Към всяка от седемте задачи (всяка задача е конкретна ситуация) има по един верен отговор, който трябва да бъде избран от девет опции (кратки текстове от рекламни рубрики от интернет източници), като всяка опция може да се използва само веднъж като верен отговор, т.е. опциите за избор към всяка задача са много повече от стандартните 3 или 4 за класическите задачи с избираем отговор. Предвидените опции за отговор са повече от задачите и са текстове, отнасящи се към една обща тема („дейности през свободното време“). Те са концептуално взаимосвързани и при съпоставянето помежду им, както и при съотнасянето им към задачите, те представляват различно атрактивни алтернативи на верните отговори. При този тестови формат дистракторите към задачите се отличават по своя относителен брой. След избор на една от опциите като верен отговор за определена задача по презумпция тази опция отпада като избор за следващите задачи. Затова вариантите за отговор при *задачите за съотнасяне* нямат еднакво голяма вероятност за избор, съответно и тяхната дискриминативна сила не може да бъде оценявана по критериите, отнасящи се за класическите задачи с избираем отговор. Статистическите данни за дискриминативната сила на дистракторите показват следното: към задача 14. има три ефективни дистрактора (В, С, D); към задача 15. има четири ефективни дистрактора (А; В; F; I); към задача 16. има шест ефективни дистрактора (С; D; G; H; I; J); към задача 17. има пет ефективни дистрактора (D; F; G; H; I); към задача 18. има пет ефективни

дистрактора (А; В; С; F; Н); към задача 19. има шест ефективни дистрактора (А; D; F; G; Н; I) и към задача 20. има пет ефективни дистрактора (В; С; D; F; Н). Ниските коефициенти на дискриминация невинаги означават, че задачите не са добри. По-скоро те са правило, отколкото изключение, защото тестовете целят проверка на различни когнитивни нива, на различни аспекти или типове постижения (Стоименова, 2000).

Асоциирането на задачите с нивата на постижения дава информация за това кои задачи са типични за групите ученици според техните нива на постижения. Трудността на задачите по групи постижения е представена на фигура № 4. Могат да бъдат направени следните изводи:

- Учениците от групата с постижения „под средното“ са били затруднени при решаването на *задачите с множествен избор* от част 2. и част 3. За тях са били трудни и преобладаващата част от *задачите за съотнасяне*.
- Над 50% от учениците от групата със „средни“ резултати са се справили успешно с почти половината от задачите и са срещнали трудности при четири *задачи с множествен избор* и при две *задачи за съотнасяне*.
- Над 50% от учениците от групата с резултати „над средното“ са се справили успешно с всички задачи от трите различни формата.

Фигура № 4 Трудност на задачите по нива на постижения



С цел проверка на взаимовръзките между трите тестови формата, на корелационен анализ са подложени сумите на точките на форматите. Стойностите на корелационните коефициенти, представени в таблица № 1, за формат 1. (0,83) и за формат 2. (0,81) са доказателство за високата степен на свързаност между тези два

формата, с което те допринасят за хомогенността на теста. По-ниската стойност на корелационния коефициент на трети формат (0,49) с другите два формата показва, че измерваните чрез него знания и стратегии за четене с разбиране се различават от измерваните знания и умения чрез другите два формата. С това той нарушава хомогенността на теста.

Формати	Корелационен коефициент	Коефициент на детерминация (общ дял вариация)	Алфа на останалите
Формат 1.	0,829	0,918	0,636
Формат 2.	0,813	0,916	0,656
Формат 3.	0,485	0,240	0,978

Таблица № 1 Статистики на форматите със сумата от другите формати

Стойностите за *Алфа* на останалите показват, че ако се изключи първият формат, надеждността ще стане 0,64, което е допустима стойност. При изключване на втория формат надеждността ще бъде 0,66, което също е допустимо. Ако се изключи третият формат, надеждността ще се покачи на 0,98. Тези стойности потвърждават общоприетото мнение в тестологията, че включването на еднотипни тестови формати на отговор в тестовете, т.е. създаването на хомогенни тестове, е предпоставка за тяхната високата надеждност. От доказателството за нарушената хомогенност на теста може да се направи заключение, че третият формат измерва различни аспекти на проверяваното умение *четене с разбиране* в сравнение с първите два формата.

Информацията за *коефициента на детерминация* на изследваните тестови формати показва, че скалите на формат 1. и формат 2. са съгласувани помежду си. Общият дял вариация при формат 1. и формат 2. е почти 92%, което се определя като много висока положителна корелация според емпиричните правила за интерпретация на корелационния коефициент (Калинов, 2010). Следователно задачите с алтернативен отговор и задачите с множествен избор имат много високи общи дялове припокриваща се информация и едва по 8% самостоятелен дял. Прието е, че при корелация над 0,91 има наличие на ниво на еквивалентност. При формат 3. общият дял на вариация (припокриващата се информация) с другите два формата е 24%, което може да се определи като много ниска положителна корелация, но достатъчна, за да участва в един

тест с другите два тестови формата. Останалите 76% са негова самостоятелна вариация, т.е. негов самостоятелен принос към измерваното умение *четене с разбиране*.

От статистиките за корелациите между резултатите от трите формата по двойки (таблица № 2) получаваме потвърждение, че корелацията между формат 1. и формат 2. е в зоната на идентичността (0,957). Високият процент припокриваща се информация може да се дължи на това, че задачите с алтернативен отговор и задачите с множествен избор проверяват знания и умения за извличане на информация от текстове за четене, които са почти еквивалентни. Корелацията на формат 3. с формат 1. е 0,49, а с формат 2. е 0,47, което е показателно, че той измерва и други аспекти освен аспектите, които покриват другите два формата. Дискриминативната сила на трети формат е много добра за практически цели (0,485), надеждността също (0,78) и неговата обща вариация с другите два формата е 24%.

	формат 1.	формат 2.	формат 3.
формат 1.	1,000	0,957	0,490
формат 2.	0,957	1,000	0,469
формат 3.	0,490	0,469	1,000

Таблица № 2 *Корелации между резултатите от трите формата по двойки*

От анализа на корелациите между сумарните резултати по трите формата могат да бъдат направени следните изводи. Корелацията между формат 1. и формат 2. е в зоната на идентичността (0,957), което е доказателство, че ако в един тест се използва комбинация от алтернативни задачи и задачи с множествен избор, ще се получи много припокриваща се информация за измерваните с теста знания и умения. Формат 3. (задачите за съотнасяне) има общ дял на вариация с другите два формата 24% и самостоятелен дял на вариация 76%, от което може да се направи заключение, че чрез включването на *задачите за съотнасяне* в теста се получава информация за знания и умения за прилагане на стратегии за селективно и детайлно четене с цел да бъдат извадени главни информации или важни подробности от кратки текстове, която не може да се получи чрез задачите от другите два тестови формата. Направените изводи от анализа на резултатите потвърждават хипотезата, че включването в структурата на тестовете на различни тестови формати с избираем отговор и различни видове текстове, с които се проверяват различни стилове и стратегии на четене, дава възможност за проверка на различните аспекти на измерваното умение и за обосновани изводи относно проверяваните умения за четене с разбиране.

Поради липсата на нормалност на разпределенията на задачите в трите формата (формат 1. е с пет задачи, формат 2. и формат 3. са с по седем задачи) е използван *стандартизираният бал* като метод за тяхното сравняване. За целта на изследването са използвани Т-скалата (скалата на Търстоун) и непараметричният тест за сравняване на резултати на Уилкоксън (Wilcoxonsigned-ranktest). Т-стойностите показват, че средната стойност на първия формат е 50,6, а средната стойност на втори формат е 50,1. Може да се направи извод, че формат 2. е незначително по-труден от формат 1. При третия формат средната стойност е 49,9. Следователно задачите от формат 3. са малко по-трудни в сравнение със задачите от другите два формата. Чрез непараметричния тест за сравняване на резултати на Уилкоксън е направена проверка като последователно са сравнени резултатите по двойки формати. Стандартизираните резултати на формат 2. и формат 1. показват, че вторият формат е само в 3 случая с по-нисък резултат отколкото първия формат. Статистиките, представени в таблица № 3, дават информация, че Z-оценката е -0,029 при препоръчан праг от 1,96 и повече. Правдоподобността¹⁶ да се получи такава и по-голяма разлика между трите ранга случайно, е в 97,7% от случаите. Може да се твърди, че няма значима разлика в резултатите в теста от задачите във формат 1. и формат 2. Сравнението на стандартизираните резултати на формат 3. със стандартизираните резултати на формат 1. показва, че разлика между тях има само в 5 случая, в които формат 3. > формат 1. Тази разлика не е статистически значима, защото Z-оценката е 0,442 при препоръчан праг от 1,96 и повече. Понеже стойността е по-малка от приетия праг, може да се направи извод, че разликата между резултатите от първи и трети формат е статистически незначима. При сравнението на стандартизираните резултати формат 2. и формат 3. е установена разлика от 9 случая с по-високи резултати по формат 2. отколкото по формат 3. Z-оценката е -0,29 при препоръчан праг от 1,96 и повече при ниво на значимост 77,2%. Следователно разликата в постиженията на учениците по формат 2. и формат 3. е статистически незначима.

Статистики	формат 2 – формат 1	формат 3 – формат 1	формат 3 – формат 2
z-оценка	-0,029	-0,442	-0,290
ниво на значимост	0,977	0,658	0,772

Таблица № 3 *Статистики на теста*

¹⁶ Виж статистическия критерий „ниво на значимост“.

От сравняването на T-стойностите на трите формата и анализа на резултатите от непараметричния тест на Уилкоксън може да се направи извод, че не се наблюдава статистически значима разлика между резултатите на учениците по трите тестови формата с избираем отговор, организирани в самостоятелни части (субтестове). Този извод потвърждава хипотезата, че отразлични тестови форматис избираем отговор, разработени чрез прилагане на правилата за създаване на задачи с избираем отговор и критериите за тяхната експертна оценка, могат да се съставят субтестове с еднаква трудност за практически цели при оценяване на уменията *четене с разбиране* на ученици от едно и също езиково ниво.

3.4. Изводи от експерименталната проверка

Анализът на статистическите данни от апробирания тест показва, че стриктното спазване на принципите и правилата за разработване на тестови задачи и критериите за тяхната експертна оценка може да допринесе за създаването на валидни и надеждни тестове за оценка на уменията *четене с разбиране*. От анализа могат да бъдат направени изводи, че няма статистически значими разлики в резултатите на учениците по трите тестови формата с избираем отговор и че чрез включването в тестове на формата *задачи за съответствие* от тестовите резултати може да се получи информация за способностите на учениците да използват различни стилове и стратегии на четене с цел извличане на информация от кратки текстове, която не може да се получи чрез задачите от вида *вярно-грешно* и *задачи за множествен избор*. Създаването на хомогенни тестове, чрез включването единствено на задачи от вида *множествен избор*, е предпочитана традиция в тестологията поради факта, че хомогенните тестове почти винаги са с висока надеждност. Тестовите за проверка на езикови компетентности са широко дефинирани, защото включват измерване на различни комуникативни компетентности, следователно те трябва да са хетерогенни. При хетерогенните тестове неизбежно някои от стойностите на дискриминативната сила на задачите могат да бъдат по-ниски, тъй като включените различни видове тестови задачи измерват широк спектър на компетентности. За да се постигне в такъв случай достатъчно висока надеждност и едновременно с това валидност, трябва да се използват видове (съдържателно) различаващи се задачи. Включването на различните тестови формати в един тест зависи от конкретните цели на теста, от определените за проверка комуникативни умения (рецептивни стилове и стратегии) и подходящите за постигане на целите на оценяването видове текстове. За да бъде един тест по-

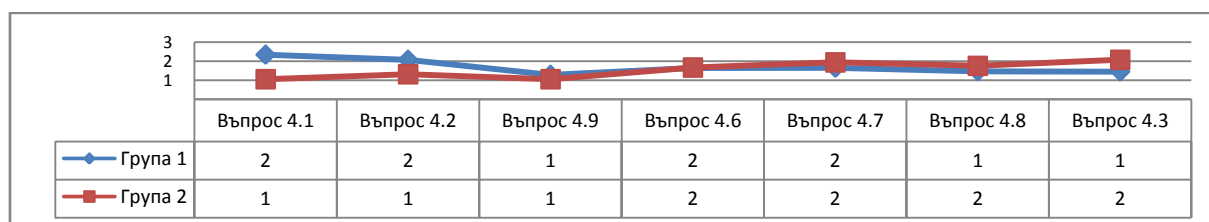
привлекателен, да покрива по-голям обем от информация за проверка, да мери различни аспекти на проверяваните умения, е необходимо да бъдат включени различни видове задачи. Всеки субтест представлява различна част от общия тест и колкото повече различни части съдържа субтеста, толкова по-точна и изчерпателна ще бъде информацията за измерваните езикови компетентности.

3.5. Представяне на анкетата и анализ на резултатите

Анкетната карта съдържа четири части. В първата част въпросите са насочени към обща информация за учениците като възраст, пол и език, на който се води всекидневната комуникация. Във втората част въпросите са свързани с опита по немски език и предишния опит в изучаването на други чужди езици, а в третата част въпросите се фокусират върху изучаването на немски език. Въпросите в четвъртата част целят събирането на информация за мотивацията за учене и нагласите към оценяването.

Средните стойности на резултатите от въпросите, свързани с увереността при решаването на задачи, показват, че учениците се чувстват уверени. Данните за честотите на отговорите на учениците по отношение на изследваното в експерименталната проверка умение за *четене с разбиране* показват следното: 74,9% от учениците са „напълно уверени” и „уверени”; „недостатъчно уверени” са 23,2% и „изобщо не са уверени” само 1,9% от учениците. Статистиките за средните стойности на въпросите, изследващи отношението на учениците към оценяването, не дадоха категоричен отговор на тези въпроси. За да могат да бъдат направени категорични изводи за отношението на учениците към оценяването, беше направена проверка чрез клъстерен анализ. Използването на метода *K-meansCluster* позволява да се открият групи ученици с близки отговори по въпросите от част 4. на анкетата. В резултат на групирането на мненията се формираха две групи: от 119 ученици или 57% (група 1.) и от 88 ученици или 43% (група 2.). Резултатите са представени на фиг. № 5.

Фигура № 5 Профил на двете групи ученици



Анализът показва, че учениците от двете групи имат еднакво мнение по въпрос 4. 9 „Когато учениците учат мотивирано, се повишават учебните резултати“. На този

въпрос и двете групи са отговорили с „да“ (код 1). Двете групи не са съгласни или нямат мнение (код 2) по въпроси 4.6 и 4.7. Мнението на учениците може да бъде обобщено по следния начин: група 1. не е съгласна или няма мнение по твърденията във въпроси 4.1 и 4.2 и е съгласна с твърденията във въпроси 4.8 и 4.3. Група 2. е съгласна с твърденията във въпроси 4.1 и 4.2, а не е съгласна или няма мнение по твърденията във въпроси 4.8 и 4.3. Направените проверки не дадоха достатъчно категорични данни за потвърждаване на хипотезата: „Тестването и оценяването е мотивиращ фактор за учащите и въздейства положително върху нагласите към ученето.“ Може да се направи извод, че за изследваната група ученици проблемът за оценяването като мотивиращ фактор няма линеен характер – оценяването може да повиши, но и да понижи мотивацията. Очакването, че оценяването е мотивиращ фактор с положително въздействие върху нагласите за учене, се потвърди частично за 40% от учениците.

Интерес представляват отговорите на учениците относно разнообразието от видове текстове и тестови задачи в тестовете по немски език. Участниците в анкетата изразяват мнение, че видовете текстове, които присъстват най-често в тестовете по немски език, са изказвания на лица, съобщения и обяви, а най-често срещаните задачи са класическите задачи с избираем отговор, следвани от задачите с алтернативен отговор, задачи с кратък свободен отговор и задачите за съответствие. Задачите с разширен свободен отговор са посочени от 42% от учениците, от което може да се направи извод, че задачите за проверка на продуктивни умения не са типични за тестовете, които учениците от анкетиранията група решават.

В анкетата бяха включени два въпроса със свободен отговор с цел събиране на информация за интереса на учениците към различните видове тестови задачи и за видовете задачи, чието решаване ги затруднява. От сравняването на резултатите от тези два въпроса може да се направи извод, че за изследваната група ученици не представляват интерес нито тестовите задачи, които ги затрудняват, нито задачите, при които не срещат затруднение. Неизбежно възниква въпросът как би могло да въздейства мотивиращо върху нагласата за учене оценяване, което се извършва чрез тестови задачи, които не са интересни на преобладаващата част от учениците. За да се отговори на този въпрос, са необходими допълнителни изследвания. Бъдещите изследвания могат да бъдат насочени към проучване на качествата на тестове, които са включени в учебните комплекси и в помагалата с тестови задачи за самооценка.

Заклучение

От проучената литература са направени изводи за характеристиките и функциите на оценяването, представени са видовете оценяване, критериите за качество на езиковите тестове и са систематизирани тестовите формати. Представянето на принципите на комуникативните тестове за оценяване на езиковите компетентности и очертаването на методите за тяхното разработване са подчинени на идеята за пълноценно използване на възможностите на тестовете в практиката. С цел оптимизиране на работата по създаване на тестове за оценяване на езикови умения са систематизирани правила за създаване на тестови задачи с избираем отговор и са изведени критерии за експертна оценка на текстове и тестови задачи за проверка на уменията *четене с разбиране*, предназначени за практическата работа на преподаватели, автори на тестови задачи и експерти. Представянето на съвременните изследвания в областта на тестването на езикови компетентности показва високата стойност на тестовете, основаваща се на обективността и точността на измерванията, които те предоставят. В комбинация с другите средства и форми на оценяване те предоставят информация за многостранна оценка.

Планираното и проведено изследване на различни тестови формати с избираем отговор за проверка на уменията *четене с разбиране* в VIII клас в средните общообразователни училища с интензивно изучаване на немски език постигна целите. От анализа на статистиките на апробираните задачи бяха направени изводи, които потвърдиха две от формулираните хипотези. Изследването на психометричните характеристики на апробираните тестови задачи потвърди първата хипотеза, че включването в структурата на тестовете на различни тестови формати с избираем отговор и на различни видове текстове, с които се проверяват различни стилове и стратегии на четене, дава възможност за проверка на различните аспекти на измерваното умение и за обосновани изводи относно проверяваните умения за *четене с разбиране*. Направените статистически проверки при изследването на трудността на задачите предоставиха категорични данни за потвърждаване на втората хипотеза, че от различни тестови формати с избираем отговор за проверка на уменията *четене с разбиране*, разработени чрез прилагане на правилата за създаване на задачи с избираем отговор и критериите за тяхната експертна оценка, могат да се съставят субтестове с еднаква трудност за ученици на едно и също езиково ниво.

Анализът на данните от анкетното проучване относно нагласите на учениците към оценяването даде отговор на поставената изследователска задача. Очакването, че оценяването е мотивиращ фактор с положително въздействие върху нагласите за учене, се потвърди частично. Направените проверки не дадоха категорични данни за потвърждаване на хипотезата: „Тестването и оценяването е мотивиращ фактор за учащите и въздейства положително върху нагласите към ученето.” Поради липсата на статистически значими разлики беше направен извод, че за изследваната група ученици проблемът за оценяването като мотивиращ фактор няма линеен характер – оценяването може да повиши, но и да понижи мотивацията. Резултатите от анкетното проучване дават основание да се формулират бъдещи изследователски задачи, насочени към допълнителни изследвания на факторите, оказващи влияние върху нагласите към оценяването на езиковите компетентности с тестове. Необходимостта от бъдещи изследвания на проблемите за разработването и използването на комуникативните тестове се обуславя от важното значение на оценяването, защото резултатите от него се отнасят до всички участници в процеса на обучението и въз основа на тях се правят изводи и се вземат решения.

Приноси

Дисертационният труд има следните приносни моменти:

- Систематизирани са предимства и недостатъци на тестовите формати с избираем отговор, на полуотворените формати на отговор и на форматите със свободен отговор.
- Изведени са принципите за създаване на комуникативни тестове за проверка на уменията *четене с разбиране*.
- Систематизирани са правилата за създаване на задачи с избираем отговор.
- Разработени са критерии за експертна оценка на текстове и тестови задачи за проверка на уменията *четене с разбиране*, предназначени за преподаватели, автори на тестови задачи и експерти.
- Установена е еднаква степен на трудност на субтестове, съставени от различни тестови форматис избираем отговор за проверка на уменията *четене с разбиране*, разработени чрез прилагане на правилата за създаване на задачи с избираем отговор и критериите за тяхната експертна оценка.

Библиография

на кирилица:

Банков К. (2012). Увод в тестологията, София: Изкуства.

Бижков, Г. (1996). Теория и методика на дидактическите тестове. второ допълнено издание.София: Просвета.

Бижков, Г., Краевски, В. (2002). Методология и методи на педагогическите изследвания. София: Университетско издателство „Св. Кл. Охридски”.

Василева Н. (2013). Българските ученици и чуждите езици: резултати от участието на България в Европейското изследване на чуждоезиковите компетентности – ESLC. *Стратегии на образователната и научната политика*, 2013, 2, 271.

Димитрова-Гюзелева, С. (1999). Оценяване на аргументативен тип есе в английския и българския език. Дисертационен труд за присъждане на научната и образователна степен „доктор“, Нов български университет.

Джалев, Л. (2013). Сравнителен анализ на приложимостта на две психометрични теории на тестовете (върху данни от теста за общообразователна подготовка). Дисертационен труд за присъждане на научната и образователна степен „доктор“, Нов български университет.

Калинов, К. (2001). Статистически методи в поведенческите и социалните науки. София: Издателство на НБУ.

Петрова, С., Н. Василева. (2013). Предизвикателства пред училищното образование, Резултати от участието на България в Програмата за международно оценяване на учениците, МОН, ЦКОКУО, София.

Стефанова, П. (2015). Методика на обучението по съвременни езици. София: Издателство на НБУ.

Стоименова, Е. (2000). Измерителни качества на тестове. София: Институт по математика и информатика, БАН.

Стойчева, Д. (1998). Проверка на учебните постижения в обучението по чужд език – значение и функции, съдържание, видове и форми, в: сп. *Чуждоезиково обучение*, кн. 5, 58-65.

Стойчева, Д. (1999). Проверка на учебните постижения по чужд език – изисквания, на които трябва да отговарят езиковите тестове и изпити. Типология на тестовите задачи. в: сп. *Чуждоезиково обучение*, кн. 4, 33-46.

Стойчева, Д. (1999а). Проверка на уменията за четене, в: сп. *Чуждоезиково обучение*, кн. 6, 24-38.

Стойнова, Ф. (1996). Тестология за учители. София: изд. „Атика“.

Трим и др. (2006). Обща европейска езикова рамка – учене, преподаване, оценяване. Варна: изд. „Релакса“.

Шопов, Т. (2013). Педагогика на езика. София: Университетско издателство „Св. Климент Охридски“.

на латиница:

Albers, H.-G., Bolton, S. (1995). Testen und Prüfen in der Grundstufe: Einstufungstests und Sprachstandprüfungen. Fernstudieneinheit 7. Berlin, München.

Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.

Arbinger, R., Herbert, S., Eberhard, T. (1977). Motivationsforschung im Bereich der Schule. Motivation. Eine Einführung in Probleme, Ergebnisse und Anwendungen. Hrsg. Eberhard Todt. Heidelberg: Quelle und Meyer.

Bachmann, L., Palmer, A. (1996). *Language testing in practice. Designing and developing useful language tests*. Oxford: Oxford University Press.

Bimmel, P. (2006). Die Not mit den Noten. In: *Fremdsprache Deutsch*, 22 (34), 34.

Bolton S. (1996). Probleme der Leistungsmessung. Lernfortschrittstests in der Grundstufe. Fernstudieneinheit 10, Langenscheidt, Goethe-Institut München.

Bolton, S., Perlmann-Balme, M. (2006). Schulische Abschlussprüfungen konzipieren – wie macht man das? Ein Werkstattgespräch zum Thema. In: *Fremdsprache Deutsch*. 22 (34), 58.

Butzkamm, W., Plum, A. (2006). Nicht für das Leben, für die Prüfung lernen wir. Der Einfluss zentraler schulischer Abschlusstests auf den Unterricht – ein Beispiel: Deutschunterricht in England. In: *Fremdsprache Deutsch*, 22 (34), 35.

Bühner, M. (2011). Einführung in die Test- und Fragebogenkonstruktion. Pearson Studium, München.

Blaska, A., Krekeler, C. (2009). Sprachtests. Leistungsbeurteilungen im Fremdsprachenunterricht evaluieren und verbessern. Baltmannsweiler: Schneider-Verlag, Hohengehren.

- Düwell, H. (1979). Fremdsprachenunterricht im Schülerurteil. Untersuchungen zu Motivation, Einstellungen und Interessen von Schülern im Fremdsprachenunterricht. Schwerpunkt Französisch. Tübingen: Narr.
- Edelhoff, C. (Hrsg.) (1985). Authentische Texte im Deutschunterricht. Einführung und Unterrichtsmodelle. München.
- Glaboniat, M. (1998). Kommunikatives Testen im Bereich Deutsch als Fremdsprache. Eine Untersuchung am Beispiel des Österreichischen Sprachdiploms. Studien-Verlag, Innsbruck-Wien.
- Glaboniat, M., Müller, M. (2006). Note „Sehr gut! – Aber in Bezug worauf? Referenzrahmen und Profile Deutsch in ihren Auswirkungen auf Prüfungen und Tests. In: *Fremdsprache Deutsch*, 22 (34), 14–21.
- Glaboniat, M. (2010). Internationale DaF-Zertifikate. Ein Überblick. In: *Deutsch als Fremdsprache*. 47 (3), 131-142.
- Grotjahn, R. (2000). Testtheorie: Grundzüge und Anwendungen in der Praxis. In: Wolff, A., Tanzer, H. (Hrsg.): Sprache – Kultur – Politik. Regensburg.
- Grotjahn, R. (2000a). Leistungsmessung und Leistungsbeurteilung. Studieneinheit, Kapitel 1. Grundlagen, Hellenic Open University, Patras, Postgraduiertenstudium Deutsch als Fremdsprache.
- Grotjahn, R. (2000b). „Determinanten der Schwierigkeit von Leseverstehensaufgaben: Theoretische Grundlagen und Konsequenzen für die Entwicklung des TESTDAF“. In: Sibylle Bolton (Hrsg.), TESTDAF: Grundlagen für die Entwicklung eines neuen Sprachtests. Beiträge aus einem Expertenseminar. Köln: VUB Gilde 2000, 7-55.
- Grotjahn, R. (2003). Leistungsmessung und Leistungsbewertung. Fernstudienbrief für den Weiterbildungs-Masterstudiengang „Deutschlandstudien.“ Schwerpunkt: Deutsche Sprache und ihre Vermittlung. FernUniversität– Gesamthochschule in Hagen.
- Grotjahn, R. (2006). Prüfen – Testen – Bewerten. In Jung, Udo O. H. (Hrsg.): Praktische Handreichung für Fremdsprachenlehrer. 4., völlig neu bearbeitete Auflage. Frankfurt/M.: Peter Lang Verlag.
- Grotjahn, R. (2008). Tests und Testaufgaben: Merkmale und Gütekriterien. In: Bernd Tesch, Eynar Leupold, Olaf Köller (Hrsg.), Bildungsstandards Französisch: konkret. Berlin: Cornelsen Verlag Scriptor 2008, 149-186.
- Grotjahn, R., Kleppin K. (2015). Prüfen, Testen, Evaluieren. Goethe-Institut, München, Klett-Langenscheidt.

- Ingenkamp, K. (1997). Lehrbuch der Pädagogischen Diagnostik: Studienausgabe. 4. Auflage. Weinheim: Beltz Verlag.
- Kniffka, G. (2003). Prüfen und Bewerten. In: Bausch K.-R., Christ H., Krumm, H.-J. (Herausgeber): Handbuch Fremdsprachenunterricht, A. Francke Verlag, Tübingen und Basel.
- Lienert, G., Raatz, U. (1998). Testaufbau und Testanalyse, Weinheim: Beltz, Psychologie VerlagsUnion.
- Mendez, C. (2013). Fehlkonstruierte Aufgaben. Häufige Fehler bei der Erstellung von Klassenarbeiten. In: *PRAXIS Fremdsprachenunterricht*. 4, 13-14.
- Mendez, C., Münchbach-Wellner, G. (2013). Respizienz = mehr als Punkte zählen. Was die Fachleitung zur Diagnose des Lernstandes und zur Förderung schwacher Lerner beitragen kann. In: *PRAXIS Fremdsprachenunterricht Französisch*. 3, 7-9.
- Neuner, G., Hufeisen, B., Kursisa, A., Marx, N., Koithan, U., Erlenwein, W. (2009). Deutsch als zweite Fremdsprache. Fernstudieneinheit 26. Berlin, München, Wien, Zürich, New York: Langenscheidt.
- Nissen, R. (2003). Formen der Leistungsmessung in der unterrichtlichen Praxis. In: K.-R. Bausch, H. Christ, H.-J. Krumm (Hrsg.): Handbuch Fremdsprachenunterricht. Tübingen und Basel: A. Francke Verlag.
- Perlmann-Balme, M. (2001). Formen und Funktionen von Leistungsmessung und -kontrolle. In: Helbig, G., Götze, L., Henrici, G., Krumm, H.-J. (Hgg.): Deutsch als Fremdsprache. Ein internationales Handbuch. 1. Halbband. Berlin New York: De Gruyter, 994-1005.
- Perlmann-Balme, M. (2006). „Das alles kann ich schon!“ Kompetenzen testen, prüfen, zertifizieren. In: *Fremdsprache Deutsch*, 22 (34), 5–13.
- Perlmann-Balme, M. (2010). Testen und Prüfen von Sprachkenntnissen. In: Krumm, H.-J., Fandrych, C., Hufeisen, B., Riemer, C. (Hrsg.), Deutsch als Fremdsprache. Ein internationales Handbuch. 2. Halbband. De Gruyter.
- Projektgruppe TestDaF (2000). „TestDaF: Konzeption, Stand der Entwicklung, Perspektiven“. *Zeitschrift für Fremdsprachenforschung*. 11 (1), 63–82.
- Rost, J. (2004). Lehrbuch. Testtheorie – Testkonstruktion. Zweite, vollständig überarbeitete und erweiterte Auflage. Verlag Hans Huber. Bern, Göttingen, Toronto, Seattle.
- Sanders, J. R. (Hrsg.) (2000). Handbuch der Evaluationsstandards: die Standards des „Joint Committee on Standards for Educational Evaluation“. 2. Auflage. Opladen: Leske und Budrich.

Solmecke, G. (Hrsg.) (1983). Motivation und Motivieren im Fremdsprachenunterricht. Paderborn: Schöningh.

Spolski, B. (1995). Measured Words. The Development of Objective Language Testing. Oxford, Oxford University Press.

Storch, G. (2009). Deutsch als Fremdsprache. Eine Didaktik. Theoretische Grundlagen und praktische Unterrichtsgestaltung. Paderborn: Wilhelm Fink GmbH & Co Verlags-KG.

Vollmer, H. J. (2003). Leistungsmessung, Lernerfolgskontrolle, Selbstbeurteilung: Überblick. In: K.-R. Bausch, H. Christ, H.-J. Krumm (Herausgeber): Handbuch Fremdsprachenunterricht, A. Francke Verlag, Tübingen und Basel.

ДОКУМЕНТИ:

Europäisches Sprachenportfolio für Jugendliche und Erwachsene (2001). Hrsg.: EDK Bern: Berner Lehrmittel- und Medienverlag BLMV.

EUROPARAT/ALTE (2012). Handbuch zur Entwicklung und Durchführung von Sprachtests. Zur Verwendung mit dem GER. Erstellt von ALTE im Auftrag des Europarats / Abteilung von Sprachenpolitik. Frankfurt am Main: telc GmbH.

Евридика (2009). Национални тестове за оценяване на учениците в Европа: цели, организация и използване на резултатите, Доклад на изпълнителна агенция “Образование, аудиовизия и култура”, Достъпен на: http://eacea.ec.europa.eu/education/Eurydice/documents/thematic_reports/109BG.pdf [Наличен на: 06.06.2016 г.]

Държавни образователни изисквания за учебно съдържание. НАРЕДБА №2 от 18.05.2000 г. за учебното съдържание. (обн., ДВ, бр. 48 от 13.06.2000 г.) Достъпен на: <http://www.mon.bg/?go=page&pageId=7&subpageId=59> [Наличен на: 05.06.2016 г.]

Учебни програми по чужди езици за VIII клас (интензивно изучаване). Достъпни на: <http://www.mon.bg/?go=page&pageId=1&subpageId=28> [Наличен на: 15.06.2016 г.]

Публикации в интернет:

Достъпът до интернет страниците е проверен за последен път на 15.06.2016 година.

Петрова, С. (2010). Училище за утрешния ден. Резултати от участието на България в Програмата за международно оценяване на учениците – PISA 2009. С., ЦКОКО 2010. Публикацията е достъпна на: <http://www.ckoko.bg/images/stories/book-2010.pdf>

Петрова, С. (2013). Предизвикателства пред училищното образование. Резултати от участието на България в PISA 2012. С., ЦКОКУО, 2013. Публикацията е достъпна на: http://ckoko.bg/upload/docs/2013-12/PISA_2012.pdf

Попандонова, Е. (2009). Педагогически измерения на мотивацията за изучаване на чужд език. Научни трудове на Русенския университет. 2009, том 48, серия 6.2. Публикацията е достъпна на: <http://conf.uni-ruse.bg/bg/docs/cp09/6.2/6.2-36-n.pdf>

Arras, U. (2006). Testen und Beurteilendes Leseverstehens in der Fremdsprache. In: *Babylonia*, 3-4/2006, 81-86. Публикацията е достъпна на: https://www.testdaf.de/fileadmin/Redakteur/PDF/Forschung-Publikationen/Arras_Beitrag_ALEG2009_Cordoba_2014.pdf

Textmapping als Methode zur Erstellung von Testaufgaben. Публикацията е достъпна на: http://www.kmk-format.de/material/Fremdsprachen/4-4-7_Textmapping_fuer_Erstellung_von_Leseaufgaben.pdf

Списък с публикациите, свързани с темата на дисертацията

1. Гергова Д., Deutsch imBlickpunkt – Компактдиск за учителя с тестови задачи по немски език за 5. клас, София, Просвета, 2006.
2. Пенева В., Д.Гергова, LerneEuropakennen! – тестови задачи по немски език за 5. клас, София, Просвета, 2006.
3. Пенева В., Гергова Д., EntdeckeDeutschland! – тестови задачи по немски език за 7. клас, София, Просвета, 2008.
4. Гергова Д., Националното външно оценяване по чужди езици – традиции и предизвикателства, *Стратегии на образователната и научната политика*, 2015, кн. 4, с. 435-449.
5. Гергова Д., Функции на тестването и оценяването по чужди езици, *Стратегии на образователната и научната политика*, 2015, кн. 6, с. 660-668.
6. Gergova D., Competence orientation in foreign language teaching and assessment – topical issues, In the Compendium, „TRADITIONS AND INNOVATIONS IN THE EDUCATION OF THE BALKAN COUNTRIES”, Compendium of the 17th International BASOPED Conference, Varna, 24.–26. September 2015.

Декларация за авторство

Долуподписаната Димитрина Иванова Гергова декларирам, че представената докторска теза на тема: *Оценяването по чужди езици в българското училище* представлява моя авторска разработка и по никакъв начин не накърнява чужди авторски права.

Подпис: