

STANDARDIZACE VÝSTUPNÍCH JAZYKOVÝCH TESTŮ C1 OBCHODNÍ ANGLIČTINY

Blanka Pojslová

Masarykova univerzita
Ekonomicko-správní fakulta
Centrum jazykového vzdělávání, oddělení na ESF
Dobrovského 767, 664 01 Bílovice nad Svitavou, Česká republika
pojslova@econ.muni.cz

Abstrakt

Centrum jazykového vzdělávání MU (CJV MU) v souladu se svou koncepcí výuky jazyků zahájilo v roce 2012 proces inovace a standardizace testování a hodnocení jazykových dovedností studentů nefilologických oborů MU. Proces inovace a standardizace testování probíhá ve dvou rovinách. Tou první je standardizace formátu výstupních jazykových testů na úrovních B1, B2 a C1 pro všechny jazyky vyučované CJV MU, což mimo jiné znamená, že se kromě tradiční gramaticko-lexikální části testují všechny čtyři jazykové dovednosti a to napříč úrovněmi i jazyky. Druhá rovina standardizace spočívá v implementaci dobré praxe vývoje jazykového testu tak, jak ji například uvádí Manuál pro vývoj jazykového testu a zkoušení (*Manual for Language Test Development and Examining*) vypracovaného asociací ALTE z podnětu Divize jazykové politiky Rady Evropy. V kontextu tohoto procesu bude představena inovace a standardizace jazykových testů obchodní angličtiny na úrovni C1 Rámce z Ekonomicko-správní fakulty MU. S použitím výsledků pretestace takto inovovaných testů bude analyzováno, jaký dopad z hlediska očekávané úspěšnosti budou mít inovované testy na testovou populaci a jak inovované testy subjektivně vzhledem ke svým potřebám vnímají sami studenti.

Úvod

V roce 2001 byla publikována anglická a francouzská verze Společného evropského referenčního rámce pro jazyky. Od svého vzniku vnesl Rámec do jazykového vzdělávání a hodnocení mnoho změn, o čemž svědčí i fakt, že Rámec dnes existuje již ve více než 30 jazycích a jeho šest úrovní (*proficiency*) dnes běžně používají jako metriku úrovně ovládnutí jazyka nejen významné testovací organizace v Evropě, ale i řada vzdělávacích institucí včetně vysokých škol.

Centrum jazykového vzdělávání Masarykovy university (CJV MU) nezůstalo stranou a po přijetí koncepce výuky jazyků pod názvem *Organizace jazykového vzdělávání na Masarykově univerzitě* kolegiem rektora MU v roce 2006 začalo realizovat svůj záměr standardizovat úroveň jazykového vzdělávání na MU a zajistit srovnatelnost jeho výsledků v mezinárodním měřítku s využitím a rozpracováním Společného evropského referenčního rámce pro jazyky (SERRJ) na akademický a odborný cizí jazyk.

Zavádění této koncepce do praxe se v první řadě dotklo sylabů kurzů, které CJV prostřednictvím svých osmi fakultních pracovišť na Masarykově univerzitě poskytuje. S tím úzce souvisí i změna přístupu ke způsobu hodnocení úrovně ovládnutí jazyka studenty MU. V rámci několika rozvojových projektů byly vyvíjeny modelové testy na úrovních B1, B2 a C1 pro akademický a odborný jazyk s cílem ověřit, jaké praktické dopady bude mít zavedení standardního cyklu vývoje jazykového testu v akademickém prostředí. Zkušenosti s vývojem modelových výstupních testů v souladu se základními testologickými principy nabyté během těchto projektů se pak staly východiskem pro projekt IMPACT, jehož jedna klíčová aktivita

se zaměřuje na inovaci a standardizaci hodnocení úrovně ovládnání jazyka studenty MU v oblasti akademického a odborného jazyka.

1 Kontext standardizace výstupních testů C1 obchodní angličtiny

1.1 Koncepce standardizace hodnocení jazykových dovedností – projekt IMPACT

Standardizací hodnocení jazykových dovedností se v kontextu testové situace CJV rozumí jednak sjednocení formátu výstupních jazykových testů a jednak implementace dobré praxe vývoje jazykového testu v souladu s *Manuálem pro vývoj jazykového testu a zkoušení*. V tradičním pojetí se u výstupních jazykových testů na CJV kladl důraz především na testování odborné a akademické slovní zásoby a gramatických struktur. Začlenění testování poslechu, čtení, psaní do výstupních testů bylo plně v kompetenci jednotlivých fakultních pracovišť a vyučujících jednotlivých kurzů.

Na tuto roztržitost formátů výstupních testů a intuitivní přístup k vývoji jazykových testů reaguje právě projekt IMPACT, který si klade za cíl sjednotit formát všech výstupních jazykových testů napříč úrovněmi B1, B2 a C1 a také vyučovanými jazyky. To v praxi znamená, že se mění formát u 67 výstupních testů u kurzů, které v rámci bakalářských a magisterských studijních programů CJV poskytuje studentům MU prostřednictvím svých fakultních pracovišť.

Standardizovaný formát výstupních testů zahrnuje všechny jazykové dovednosti – poslech, čtení, psaní a mluvení a tradičně i gramaticko-lexikální část. Výstupní testy mají také jednotné 100 bodové hodnocení, což usnadní srovnatelnost testů a úrovně studentů napříč fakultami MU. Váhy jednotlivých sub-testů na celkovém bodovém hodnocení výstupního testu si dle svých specifických podmínek u jednotlivých úrovní B1, B2 a C1 zvolí jednotlivá fakultní pracoviště sama. Také testové metody, kterými se budou jednotlivé dovednosti na fakultách testovat, si pracoviště dle svých zkušeností a specifických potřeb vyberou sama.

2 Standardizace výstupních testů na Ekonomicko-správní fakultě

2.1 Charakteristika výstupní testů na Ekonomicko-správní fakultě

Jazykové oddělení na Ekonomicko-správní fakultě patří k největším oddělením CJV na MU a ve své výuce se zaměřuje především na odborný ekonomický jazyk, nicméně neopomíjí ani prvky akademického jazyka. Jazykové oddělení ESF vyvíjí výstupní testy v paralelních verzích na úrovních B1, B2 a C1 dle Rámce pro jazyk anglický, německý, francouzský a španělský. Jedná se o testy kriteriální, které na rozdíl od testů normativních hodnotí studenty čistě na základě jejich schopností a nezávisle na znalostech jejich kolegů.

Pro kriteriální testy je charakteristické, že je stanoven okruh relevantních jazykových oblastí a minimální úroveň, kterou je při testu třeba zvládnout. Typicky minimální počet bodů, kterého musí testovaný dosáhnout, aby u testu uspěl. Výstupní testy jsou dále typu *pro-achievement*, jelikož v sobě kombinují prvky testů jazykové způsobilosti (*proficiency*) – sub-testy čtení, poslech, mluvení a psaní - a testu didaktického (*achievement*) – sub-test gramaticko-lexikální část [3].

2.2 Podoba inovovaných výstupní testů na Ekonomicko-správní fakultě MU

Tvůrci výstupních testů na Ekonomicko-správní fakultě MU plně v souladu s koncepcí CJV zahrnuli do inovovaných testů všechny čtyři jazykové dovednosti (čtení, poslech, psaní a mluvení) spolu s tradiční gramaticko-lexikální částí a vývoj první sady testů prošel standardním vývojovým cyklem. Váhy jednotlivých sub-testů na celkovém 100 bodovém

hodnocení se u jednotlivých cílových úrovní liší, nicméně pro všechny vyučované jazyky jsou stejné. Na úrovni C1 je rozdělení vah mezi sub-testy rovnoměrné. U úrovní B2 a B1 je posílena gramaticko-lexikální část na 30 % na úkor poslechu a psaní, jejichž váha v testu tvoří 15 %, souhrnně viz Tabulka 1.

Tab. 1: Váhy jednotlivých sub-testů na celkovém počtu bodů inovovaného testu ne ESF.

Úroveň	Poslech	Čtení	Gram-lex	Psaní	Mluvení
B1	15 %	20 %	30 %	15 %	20 %
B2	15 %	20 %	30 %	15 %	20 %
C1	20 %	20 %	20 %	20 %	20 %
Původní C1	16,6 %		66,6 %	16,6 %	nestanoveno

Zdroj: Testové specifikace výstupních jazykových testů na CJV, ESF MU

Jednotná váha sub-testů pak při bodové dotaci 1 bodu za položku znamená, že v rámci jedné dovednosti připadá na testy pro různé jazyky stejný počet testových položek. U jedné úrovně a v rámci jednoho sub-testu se testeři snažili napříč jazyky sjednotit i testové metody. To se podařilo jen do určité míry, protože bylo třeba respektovat specifika jednotlivých jazyků a zároveň různé tradice testování v těchto jazycích.

2.4 Inovace výstupního testu úrovně C1 pro odbornou ekonomickou angličtinu

Výstupním jazykovým testem na úrovni C1 ukončují studenti bakalářských a magisterských programů prezenčního studia na ESF čtyřsemestrální kurz odborného a akademického jazyka. Jedná se o povinný předmět pro všechny studenty, kteří si zvolí angličtinu jako svůj Jazyk I. V průběhu 1. a 2. semestru je časová dotace výuky 4 hodiny, zatímco ve 3. a 4. semestru jsou to hodiny dvě.

Současný výstupní test zahrnuje šest sub-testů: poslech, psaní, slovní zásoba, překlad, gramatika a mluvení. V testu tedy úplně chybí sub-test čtení. Dovednost čtení je však do určité míry zahrnuta v sub-testech mluvení a psaní. Současný formát sub-testu mluvení zahrnuje detailní čtení ekonomického textu v rozsahu 500-600 slov a následnou diskuzi se zkoušejícím nad obsahem textu. U sub-testu psaní pak mají studenti za úkol napsat shrnutí textu v rozsahu 370-450 slov.

Výstupní test ve své současné podobě je tedy kombinací testu hodnotícím úroveň ovládnutí jazyka prostřednictvím testování jednotlivých dovedností (*skilled-based test*) a testu, u kterého testovaný musí kombinovat více dovedností, aby splnil úkol (*task-based test*) [4]. Součástí původního testu je také překlad pasáže z českého jazyka do anglického, což lze považovat za krajně nestandardní metodu testování úrovně ovládnutí jazyka, která je zároveň značně problematická z hlediska spolehlivosti hodnocení.

Při vývoji testových specifikací inovovaného výstupního testu úrovně C1 vycházeli testeři z nové koncepce hodnocení úrovně ovládnutí jazyka na CJV, přičemž na všechny sub-testy hodnotící dovedností poslech, čtení a gramaticko-lexikální část připadá 20 testových položek, což splňuje kritéria jedné z rovin standardizace výstupní testů na CJV týkající se formátu testů. Druhá, neméně důležitá rovina standardizace hodnocení úrovně zvládnutí jazyka spočívá v implementaci standardního vývojového cyklu na inovované výstupní testy.

Vývojový cyklus jazykového testu začíná vývojem testových specifikací, pokračuje tvorbou testových položek a jejich moderace. Po sestavení paralelních verzí testů následuje jejich pilotáž, kterou završí analýza výsledků pilotáže. Výstupem této analýzy je úprava testových specifikací a revize sporných testových položek [1] a [2].

2.4.1 Testové specifikace inovovaného výstupního testu C1

Testové specifikace ve své části Základní charakteristiky testu definují účel inovovaného výstupního testu na úrovni C1 *jako ověření dosažených znalostí, ověření stanovené jazykové úrovně podle SERR a ohodnocení studenta známkou*. Testovaný jazyk je pak charakterizován jako *jazyk pro specifické účely* a test je typu *pro-achievement v papírové verzi*. Testová populace je charakterizována jako *čeští a slovenští studenti, kteří absolvovali 4 semestry výuky anglického jazyka (Jazyk 1) v bakalářských a magisterských programech Ekonomicko-správní fakulty MU*. Testové výsledky se pak dle testových specifikací použijí k rozhodnutí, zda udělit či neudělit povinné kredity.

Test se skládá z pěti sub-testů a celková časová dotace písemné části testu je 120 minut, přičemž váhy jednotlivých dovedností jsou 20 %. Jednotlivé sub-testy mají být podle specifikací administrovány v pořadí poslech, psaní, gramaticko-lexikální část a čtení a úspěšné složení písemné části zkoušky podmiňuje připuštění k sub-testu mluvení. Tematické oblasti testu zahrnují *témata obchodní angličtiny z učebnic Intelligent Business Upper-Intermediate a English for Business Studies*. Minimální hranice úspěšnosti (*passmark*) u sub-testů poslech, čtení a gramaticko-lexikální část *budou stanoveny metodami doporučenými v základních dokumentech SERRJ (standard setting)*. Hranice úspěšnosti u sub-testu psaní a mluvení *budou stanoveny metodami benchmarkingu*. Testy jsou opravovány ručně bez kontroly dalším vyučujícím a *za položku se uděluje stanovený počet bodů bez možnosti přiznání částečného kreditu s výjimkou jedné testové úlohy v gramaticko-lexikální části, kde se částečný kredit připouští*.

2.4.2 Charakteristika současného výstupního testu C1

Formát současného výstupního testu není pevně zakotven v testových specifikacích a postupně se ustálil na bodovém hodnocení 120 bodů, přičemž váha jednotlivých sub-testů na celkovém počtu bodů je rozložena nerovnoměrně s tradičním důrazem na gramaticko-lexikální část. Vezmeme-li v úvahu, že při hodnocení překladu se klade největší důraz na hodnocení gramatické a lexikální správnosti převedeného textu z českého jazyka do anglického, tak celková gramaticko-lexikální část představuje 66,6 % z celkového počtu bodů a neúměrně tak zatlačuje do pozadí ostatní dovednosti, které jsou z hlediska hodnocení úrovně ovládnutí jazyka s těžší, viz Tab. 1. U testu tak převažuje jeho didaktický charakter, při kterém se hodnocení soustřeďuje na to, co bylo předmětem výuky a už méně na to, jak umí student své znalosti aplikovat ve skutečném životě.

U původního formátu testu také není přesně stanovena váha sub-testu mluvení na celkovém hodnocení. Dovednost mluvení podobně jako dovednost psaní je hodnocena ryze subjektivně na základě dojmu zkoušejícího. U dovednosti psaní zkoušející písemný výstup hodnotí body na škále 0-20 bodů, u mluvení pak na základě výkonu u ústní části zkoušky může dojít ke zlepšení či zhoršení známky z písemného testu o dva stupně.

3 Srovnání formátů současného a inovovaného výstupního testu C1

3.1 Poslech s porozuměním

Inovovaný a původní test se neliší pouze ve vahách kladených na jednotlivé dovednosti, ale také v typech testových úloh testující jednotlivé dovednosti. Dovednost porozumění slyšenému se v současném testu ověřuje prostřednictvím dvou formátů. Každý formát obsahuje dvě testové úlohy, z nichž jedna je otevřená a druhá uzavřená. U formátu prvního typu se uzavřená úloha konstruuje jako úloha s výběrem z více odpovědí (*multiple-choice questions*), zatímco otevřenou úlohu reprezentuje zjednodušené shrnutí nahrávky, do které má

student doplňovat chybějící výrazy. U formátu druhého typu uzavřenou úlohu představují tvrzení, u nichž se má testovaný rozhodnout, zda jsou pravdivé či ne. Otevřenou úlohu přestavuje úloha se stručnou odpovědí (*short-answer question*).

Tvůrci testu se snažili tímto formátem, tedy kombinací otevřené a uzavřené úlohy, vzájemně kompenzovat nevýhody uzavřených a otevřených otázek. Nicméně z testologického hlediska není přijatelné, aby se u paralelních verzí testů lišil typ testových úloh. Navíc úloha s výběrem z více odpovědí je bodována 2 body za položku, aniž by se dalo prokázat, že položka je dvakrát tak obtížná jako ostatní položky testující poslech s porozuměním. U úlohy se stručnou odpovědí jsou body pro jednotlivé položky přidělovány v rozmezí 1-3 body, aniž by bylo vždy zřejmé, čemu vyšší počet bodů odpovídá. To pak má své dopady z hlediska spolehlivosti testu, protože různí opravující přidělují body za 2-3 bodové položky podle ne zcela jasných kritérií.

Podle testových specifikací pro inovovaný sub-test poslech s porozuměním je konstrukt definován jako *porozumění slyšenému odbornému ekonomickému textu různého charakteru, se zaměřením na globální i detailní porozumění s časovou dotací 30 minut*. Při vytváření specifikací se testeři u tohoto sub-testu pokusili výše zmíněné nedostatky původního testu eliminovat. Typy testových úloh vybírali tak, aby co nejvíce snížili testování jiných dovedností – tedy psaní a čtení. Samozřejmostí je, že u paralelních verzí testu je dodržován identický formát úloh.

Při váze 20 % má sub-test poslech s porozuměním 20 testových položek, přičemž položka je hodnocena buď jako správně nebo špatně zodpovězená. Za každou položku může testovaný obdržet 0 nebo 1 bod, což vede k celkovému zvýšení konzistentnosti v opravování různými hodnotiteli a tak přispívá k celkové spolehlivosti testu.

Ve snaze co nejvíce při testování poslechu s porozuměním omezit psaní, byly zvoleny dvě uzavřené úlohy a jedna otevřená. Pět položek připadá na úlohu typu Pravda/Nepřavda, kde se na autentické nahrávce zpráv s ekonomickou tematikou testuje detailní porozumění slyšenému. Druhá uzavřená úloha je typu úlohy s výběrem ze čtyř odpovědí, u které se na rozhovoru dvou a více lidí na ekonomické téma testuje globální i detailní porozumění. Otevřenou úlohu představuje doplnění 1-3 slov do izolovaných vět shrnujících slyšené ve tvaru, ve kterém odezněla v poslechu nebo je jeho parafrázi. Textem, na kterém se detailní porozumění slyšenému testuje, je monolog na odborné téma – typicky přednáška.

3.2 Čtení s porozuměním

Jestliže testeři mohli při vývoji specifikací sub-testu poslech s porozuměním alespoň částečně využít svých zkušeností z testování této dovednosti v původním formátu testu, u čtení tuto možnost neměli, protože se tato dovednost samostatně netestovala. Zahnutí této dovednosti do sub-testu psaní a mluvení mělo funkci vstupu pro hodnocení produktivních dovedností a opět neexistovala kritéria, která by hodnotila, do jaké míry testovaný textu porozuměl a jaký to pak mělo vliv na hodnocení jeho výstupu u produktivních dovedností. Z tohoto důvodu se čtení v inovovaném formátu testu úplně z testování produktivních dovedností eliminovalo a začalo se testovat samostatně.

Čtení s porozuměním má podobně jako ostatní sub-testy váhu 20 % a 20 testových položek hodnocených 0 nebo 1 bod je rozloženo do tří typů testových úloh, na které mají testování časovou dotaci 40 minut. Konstrukt je u čtení s porozuměním definován jako *porozumění psanému textu, ověření schopnosti identifikovat souvislosti v něm, ověření pochopení koheze textu*.

První testová úloha s výběrem ze čtyř odpovědí testuje detailní porozumění psanému textu, kterým je odborný nebo populárně-naučný text v rozsahu 150-200 slov. Z nabídky 4 výrazů testovaný v jeho 8 položkách vybírá správný výraz. U druhé úlohy testování do odborného nebo populárně-naučného textu v rozsahu 250-300 slov doplňují kohezní prostředky. Mají na výběr z 8 výrazů, které doplňují do 6 mezer. Testovanou dílčí dovedností je pochopení koheze a užití kohezní prostředků. Poslední úloha je opět typu úlohy s výběrem ze čtyř odpovědí a testuje se detailní a globální porozumění psanému textu, který je dle specifikací charakterizován jako odborný nebo populárně-naučný text v rozsahu 350-450 slov. Použitím uzavřených typů úloh v sub-testu čtení s porozuměním se testeři snažili eliminovat testování jiných dovedností než čtení a to především psaní.

3.3 Gramaticko-lexikální část

Třetí sub-test inovovaného výstupního testu slučuje tři sub-testy současného testu, kterými jsou slovní zásoba, překlad a gramatika, do jednoho gramaticko-lexikálního sub-testu. Z původních čtyř testových úloh hodnotících znalost gramatických struktur přesně vymezených sylabem kurzu zůstala v inovovaném testu pouze úloha typu transformace izolované věty na jinou větu při zachování sémantického významu a změně formy (gramatického významu). Ostatní typy úloh – doplňování závislých předložek do vět, testování slovesných vazeb a úloha typu Pravda/ Nepravda s výběrem ze 4 možností – do inovovaného testu nebyly zahrnuty. Důvodem byla snaha hodnotit, jak testovaný gramatické struktury používá zejména v produktivních dovednostech než testovat jejich znalost na diskrétních položkách. Konstrukt testovaných úlohami eliminovaných z inovovaného testu je navíc součástí konstruktů v jiných sub-testech inovovaného testu.

Naproti tomu u testování slovní zásoby se z původního 4 testových úloh v inovovaném formátu sub-testu zachovaly 3 testové úlohy. Jedná se o překlad pěti sousloví o dvou sémantických částech z češtiny do angličtiny, kde je testovanou dílčí dovedností znalost odborné slovní zásoby. Odborná slovní zásoba se testuje také ve druhé testové úloze, ve které mají testovaní do pěti izolovaných vět doplnit 5 sloves z 10 nabízených. Důraz se u této testové úlohy klade na znalost kolokací v odborném jazyce. Třetí úloha – slovtvorba byla také zachována z původního sub-testu Slovní zásoba. Dílčí testovanou dovedností je znalost slovní zásoby a schopnost tvořit slova od stejného slovního základu.

Sub-test Gramaticko-lexikální část je jediným čistě didaktickým testem (*achievement test*), protože se zaměřuje na měření znalostí korespondujícím se sylabem kurzu a testovaní zde přesně vědí, jaký rozsah gramatických a lexikálních prostředků zde mohou očekávat. Z tohoto důvodu si testeři dovolili tyto prostředky testovat prostřednictvím izolovaných vět a nikoli jak by z testologického hlediska bylo případnější v kontextu [4]. Velkou roli zde také hrála praktičnost testu. Při vývoji testu v akademickém prostředí se testeři musí vyrovnávat s různými omezeními a volba izolovaných položek v této části výstupního testu přinejmenším snižuje časové nároky na vyhledávání textů, které by v požadované šíři pokryly především slovní zásobu danou sylabem kurzu.

4 Analýza evaluačního dotazníku pro studenty

4.1 Administrace evaluačního dotazníku

Pilotáž pěti verzí inovovaných výstupních testů C1 proběhla v květnu 2013 a zúčastnilo se jí 386 studentů, kteří předmět v daném semestru ukončili ještě prostřednictvím původního testu. Studenti byli tudíž dobře obeznámeni s původním formátem testu a mohli jej dobře porovnat s inovovaným formátem.

Zpětná vazba od studentů byla získána dvojím způsobem. Jednak krátkými rozhovory bezprostředně po skončení testu a jednak prostřednictvím evaluačních dotazníků pro studenty. Dotazník studenti obdrželi spolu s testem a byli vyzváni, aby se s jeho obsahem seznámili dříve, než začnou na testu pracovat. K vyplnění evaluačního dotazníku měli ale přistoupit až po dokončení testu. Dotazník sestával z 10 otázek. Devět z nich bylo polo-uzavřených a studenti na ně odpovídali ANO / NE a u každé otázky byli vyzváni k doplňujícímu komentáři. Desátá otázka byla otevřená a poskytla studentům prostor se k testu vyjádřit obecně.

Otázky z evaluačního dotazníku pro studenty lze rozdělit do 3 kategorií. První tři otázky se týkaly charakteristik testu, které jsou pro paralelní verze výstupního testu stejné. Jedná se o srozumitelnost a jednoznačnost instrukcí k jednotlivým testovým úlohám a názornost příkladů. Do druhé kategorie patří otázky č. 4-9, které se blíže dotazovaly na sub-testy poslech s porozuměním a čtení s porozuměním. Jednalo se o názor respondentů na zajímavost textů u obou sub-testů, jednoznačnost odpovědí a možnosti položku odpovědět, aniž by bylo třeba vyslechnout nebo přečíst. U sub-testu poslech se pak otázky č. 5 a č. 6 soustředily na časový prostor pro obeznámení se s úkolem a zaznamenání odpovědí. Poslední kategorii představuje desátá otázka určena pro doplňující komentář.

4.2 Diskuze – polo-uzavřené otázky

U první kategorie otázek týkající se instrukcí a příkladů, považuje instrukce za srozumitelné 98,7 % respondentů a jednoznačné 96,6 % respondentů. U první otázky týkající se srozumitelnosti instrukcí však respondenti odpovídající záporně neuvedli konkrétní příklady úloh, jejichž zadání pro ně bylo nesrozumitelné. U druhé otázky týkající se jednoznačnosti instrukcí respondenti odpovídající záporně upozorňují na tři úlohy, u nichž chybí informace o počtu správných odpovědí. Otázku týkající se názornosti příkladů zodpovědělo kladně 96,6% respondentů, nicméně dle volných odpovědí u záporně odpovídajících respondentů lze soudit, že někteří v této otázce odpovídali na první dvě otázky a bude třeba ji pro příští pilotáži přeformulovat.

Z druhé kategorie otázek vztahujících se k sub-testům poslech a čtení vyplývá, že 89,3 % respondentů shledává poslechové texty zajímavé a podobně je tomu i u textů pro čtení, které považuje za zajímavé 89,5 % respondentů. Respondenti odpovídající na tuto otázku záporně se většinou shodují na tom, že nezajímavý až monotónní je text v monologu u otevřené poslechové úlohy. Většinou se v paralelních verzích testu jedná o souvislý odborně zaměřený akademický text, nicméně obsahově se vztahující k sylabu kurzu. U textů ke čtení se studenti odpovídající na otázku týkající se jejich zajímavosti výrazně na jednom textu neshodují, nicméně z jejich volných odpovědí mohou být pro tvůrce testu a evaluačního dotazníku inspirující následující volné odpovědi.

- *Zajímavost textu mě v ostrém testu nezajímá, mým cílem je získat body, na sebeobohacení není čas*
- *Texty byly velmi zajímavé, jeden text bych možná koncipoval makroekonomicky, možná TASK 8 – tam bych zvolil jiný text, ale každému vyhovuje něco jiného*
- *TASK 8 – co je mi po chairmanovi?:) – nicméně mi to nevadí, podstatou tohoto readingu je něco jiného než zaujmout čtenáře 😊*
- *TASK 10 – tendenční feministický článek, hlavně poslední odstavec.*

Zdroj: Evaluační dotazníky z pilotáže 2013 výstupních jazykových testů na CJV, ESF MU

Fig. 1: Volné odpovědi hodnotící zajímavost textů použitých v testech C1

Z odpovědí na otázku č. 7 týkající se jednoznačnosti odpovědí u testových úloh vyplývá, že 67 % respondentů považuje odpovědi na testové otázky za jednoznačné a tedy z nabízených možností považují za správnou pouze jednu. Nicméně 30 % respondentů uvádí konkrétní testové úlohy a položky, u kterých by považovali za správnou i více než jednu odpověď. Tvůrci jednotlivých sporných položek pak tyto případy vyhodnotí a spolu s položkovou analýzou a analýzou distraktorů může tato zpětná vazba od studentů sloužit jako další vodítko k výběru položek, které bude třeba revidovat.

Obdobně tomu bude u i testových položek zmíněných 12 % respondentů, kteří si myslí, že jsou schopni na testovou otázku odpovědět na základě osobní zkušenosti či svých znalostí, aniž by museli číst nebo vyslechnout text, podle něhož mají na testovou položku odpovědět. Nicméně 86 % respondentů uvádí, že v sub-testech poslech a čtení takovéto úlohy nejsou. Bohužel dotazník je anonymní, ale bylo by jistě zajímavé zjistit, zda správné odpovědi výše zmíněných 12 % testovaných na tento typ testové otázky také koresponduje se správnou odpovědí dle klíče.

Tvůrce testů velmi zajímaly odpovědi respondentů na otázku č. 6 týkající se dostatečného množství času pro zaznamenání odpovědi v sub-testu poslech a to především u úlohy, kde měli do shrnujících vět napsat chybějící výraz či výrazy. Důvodem byla změna administrace poslechu u inovovaného testu. Původně se u sub-testu poslech dělaly při všech poslechových úlohách pauzy, aby měli testovaní dostatek času napsat nebo označit odpovědi. Testové položky byly v původním testu často konstruovány tak, že bez pauz nebylo možné relativně dlouhou odpověď vůbec napsat. U inovovaného testu proto byla zvolena forma volných odpovědí v rozsahu maximálně 3 slov a při tvorbě položek a jejich moderaci se velmi dbalo na to, aby jednotlivé položky nenásledovaly příliš rychle za sebou. Navíc praxe s pauzami není obhajitelná u uzavřených otázek a naopak mnohdy působí rušivě.

Na otázku č. 6 týkající dostatečnosti časového prostoru na zaznamenání odpovědi odpovídá 74 % respondentů kladně a 26 % respondentů záporně. Tito respondenti také ve volných odpovědích zmiňují jako problematické právě doplňování výrazů a shodně uvádí, že jim chyběly průběžné pauzy. Nicméně lze však konstatovat, že změna administrace sub-testu poslech je pro většinu testovaných schůdná a jakmile se tato změna odrazí i ve výuce, během které budou studenti systematicky připravováni na tento způsob administrace poslechu, tak lze očekávat, že procento kladných odpovědí bude vyšší.

Pátá otázka také souvisela s administrací sub-testu poslech a to s dostatečností časové dotace pro obeznámení se s úkolem, než dojde k jeho vyslechnutím. Zde 86 % respondentů uvádí, že časový prostor byl dostatečný. Zbývající 14 % respondentů, kteří odpověděli záporně, ve svých volných odpovědích jako problematickou z hlediska obeznámení považují úlohu s doplňováním výrazů a ojedinele také uzavřené úkoly Pravda/ Nepravda a MCQ. Jejich výtky se týká nedostatečnosti času na přečtení otázek a opět zmiňují absenci pauz.

4.3 Diskuze – volné odpovědi

Pro tvůrce testů byly pak velmi přínosné volné odpovědi v desáté otázce, protože zde se respondenti měli vyjádřit k testu jako celku. Z volných odpovědí vyplývá, že testování inovaci vnímají jako pozitivní změnu, která více odpovídá jejich skutečným potřebám v akademickém i profesním životě. Velmi kladně hodnotí začlenění sub-testu čtení (to také vplynulo z rozhovorů bezprostředně po skončení testu) a menší důraz na testování slovní zásoby a gramatiky.

Většinou shledávají test lehčím, ale je třeba vzít v úvahu fakt, že se k obtížnosti testu vyjádřili, aniž by věděli, jak u něj uspěli. Následné srovnání ukázalo, že úspěšnost u inovovaného testu je srovnatelná s úspěšností u současného formátu. Zde je však třeba vzít

v úvahu fakt, že se jednalo o pilotáž, kde testování obvykle podávají horší výkon než u ostrého testu.

Z volných odpovědí, které hodnotí test pozitivně, pak stojí za pozornost následující.

- *Test se mi zdá vyvážený, obtížnost srovnatelná s nynějším testem.*
- *Mnohem časově vyváženější a míň časově stresující test, než test dělaný doposud.*
- *Test se mi zdál v porovnání s minulými celkem lehký.*
- *Tento test je lepší než jeho současná forma, lepší a zajímavější poslech (TASK 3). Lepší gramatická část (bez možností abcd), lepší překlad, zajímavá část Reading.*
- *Test je vyvážený a srozumitelný, úlohy odpovídají náročnosti testu. Moje jediné výhrady směřují k poslechům, které jsou někdy nesrozumitelné i pro vysoce pokročilé – nejde o náročnost, ale o rychlost a přízvuk.*
- *Moc dlouhé. Ale cením si toho, že místo dlouhého překladu tu bylo porozumění textu, které považuji za poměrně důležité - za to určitě palec nahoru.*
- *Chybí zde nejtěžší (alespoň vždy pro mne) část testu, Grammar, doplňování sloves v různých tvarech + přípony či předložky. Z pohledu žáka dobrý krok :-)*
- *Pěkný reading (styl cvičení).*
- *Moc se mi líbí, jak je to zpracováno. Hlavně doplnění o reading (pokud není i teď na závěrečné zkoušce) je super. Byl o něco lehčí než klasický zápočtový test, méně gramatiky a slovní zásoby tomu prospělo.*
- *Test mi připadal lehčí než současné zápočtové testy (méně gramatiky a slovní zásoby). Přijde mi to jako pozitivní změna. Část reading je podle mě dnes důležitější než mít v testu jenom plno gramatiky.*
- *Tento test se mi zdál lehčí než ten, který budeme psát my.*
- *Test mne připadá vyváženější, výrazně ubylo vocabulary a grammar, možná až příliš v případě vocabulary.*
- *Zajímavá forma testu – lepší než v současnosti, člověk využije své reálné znalosti*
- *V starších testech méně překladů, potěšilo mne přidání čtení s porozuměním*
- *Angličtiny v praxi, ne založené hlavně na slovní zásobě.*
- *Výrazně zajímavější než současná forma.*
- *Formát testu se mi zdá lepší než ten současný. Nezohledňuje tolik slovní zásobu, ale je více orientován na poslech a porozumění textu. Odhalí tedy více, jak je na tom student s angličtinou celkově, nejen to, jak moc driloval slovička.*

Zdroj: *Evaluacní dotazníky z pilotáže 2013 výstupních jazykových testů na CJV, ESF MU*

Fig. 2: *Volné odpovědi hodnotící, jak testování subjektivně vnímají inovaci testů C1*

V menší míře se objevují kritické poznámky týkající se především menšího důrazu na slovní zásobu, kterou tito studenti mnohdy vidí jako zásadní. Pro slabší studenty představovala část *achievement* v současném testu tvořeném slovní zásobou, překladem a gramatikou možnost, jak se na zkoušku “naučit” a tak předmět úspěšně ukončit. Ve velké míře také zaznívaly výtky vůči náročnosti poslechů. Opět by stálo za podrobnější prozkoumání, jak tyto výtky korelují s celkovou úspěšností testovaných a jejich úrovní ovládnutí jazyka.

- Podle mě je příliš málo prostoru kladen na VOCABULARY a GRAMMAR, přitom to jsou podle mě nejdůležitější části kurzu i jazyka vůbec, zejména pak odborná terminologie je důležitá. Přidal bych víc lexikálních cvičení na úkor READINGu.
- Gramatika dostala málo prostoru na reálné zhodnocení úrovně studentů
- Druhý poslech velmi nepřehledný, monotónní.
- v porovnání se stávající zkouškou mi přijde, že je zde kladen malý důraz na slovní zásobu, na kterou se klade důraz celé 4 semestry
- Na tak obrovskou slovní zásobu mi přijde 5 slovních spojení moc málo, u Task 4
- Poslech se mi osobně zdál obtížnější než u dřívějších testů, naopak jsem ráda, že zde nebyl dlouhý překlad slova od slova

Zdroj: *Evaluací dotazníky z pilotáže 2013 výstupních jazykových testů na CJV, ESF MU*

Fig. 3: *Volné odpovědi hodnotící, jak testování subjektivně vnímají inovaci testů C1*

Závěr

Proces inovace výstupního testu C1 obchodní angličtiny, který probíhá v širším kontextu inovace a standardizace výstupních testů na CJV MU, prošel svým prvním vývojovým cyklem. Pro vývojový cyklus jazykového testu je charakteristické, že se jedná o proces iterativní. Každá další fáze tohoto procesu nutí tvůrce testů, aby se vrátili do té předcházející a přehodnotili některá svá rozhodnutí. A právě výstupy první pilotáže inovovaných testů C1 se významně odrazí v úpravách testových specifikací a jednotlivých testových položek.

Zpětná vazba od testovaných tvoří bezpochyby podstatnou součást výstupů pilotáže a její analýza ukázala, že jako celek je inovovaný test vnímán testovanými jako krok směrem k nástroji, který měří úroveň ovládnutí jazyka a nikoli jen míru zvládnutí sylabu kurzu či znalostí o jazyce. U konkrétních úloh a položek je pak třeba brát názor testovaných v potaz, protože sebelépe provedená fáze moderace nikdy neodhalí nedostatky testu tak, jako jeho pilotáž a její nedílná součást, kterou je zpětná vazba od testovaných.

Literatura

- [1] ALTE/COUNCIL OF EUROPE. *Manual for Language Test Development and Examining for use with the CEFR*. Strasbourg: Council of Europe [online]. 2011 [cit. 2013-09-10] Dostupné z: <http://www.coe.int/T/DG4/Linguistic/ManualLangageTest-Alte2011_EN.pdf>.
- [2] BACHMAN, L.F., PALMER, A.S., 1996. *Language Testing in Practice*. Oxford: OUP 1996. 377 s., ISBN 978-0-19-437148-3.
- [3] Council of Europe. *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Strasbourg: Council of Europe [online]. 2001 [cit. 2013-09-11] Dostupné z: <http://www.coe.int/t/dg4/linguistic/source/framework_en.pdf>.
- [4] CARR, T.N., 2011. *Designing and Analyzing Language Tests*. Oxford: OUP 2011. 360 s., ISBN 978-0-19-442297-0.

BUSINESS ENGLISH LANGUAGE TEST STANDARDISATION

The Language Centre (LC) of Masaryk University launched a language test and assessment standardisation and innovation process in accordance with its language policy in 2012. The standardisation process involves two distinct dimensions. The first one represents standardising the format of its language tests at the B1, B2 and C1 levels. The other scope consists in implementing a set of common standards for language test development. The standardisation and innovation of Business English tests at the C1 level of the CEFR will be presented within this framework. Using the outcomes of the pretesting phase, the impact of these tests on the target population will be analysed in terms of the expected success rate and students' subjective perception of the newly developed tests with respect to their needs.

STANDARDISIERUNG VON TESTS IM WIRTSCHAFTSENGLISCHEN

Der Fremdsprachenlehrstuhl an der Masaryk-Universität (MU) eröffnete im Jahre 2012 in Übereinstimmung mit seiner Unterrichtskonzeption den Prozess der Innovation und der Standardisierung des Testens und der Bewertung der sprachlichen Fertigkeiten der Studenten nichtphilologischer Studienfächer an der MU. Der Prozess von Innovation und Standardisierung des Testens verläuft auf zwei Ebenen. Die erste Ebene besteht in der Standardisierung des Formats der Sprachtests auf den Stufen B1, B2 und C1. Die zweite Ebene der Standardisierung beruht auf der Implementierung einer guten Entwicklungspraxis von Sprachtests. Im Kontext dieses Prozesses werden Innovation und Standardisierung von Sprachtests in Wirtschaftsenglisch auf dem Niveau C1 der Fakultät für Ökonomie und Verwaltung der MU vorgestellt. Unter Nutzung der Ergebnisse der Prätestation der so innovierten Tests wird analysiert, was für eine Auswirkung die innovierten Tests auf die getestete Population unter dem Gesichtspunkt der erwarteten Erfolgsquote haben wird und wie die innovierten Tests subjektiv von den Studenten selbst im Hinblick auf deren Bedürfnisse wahrgenommen werden.

STANDARYZACJA WYJŚCIOWYCH TESTÓW JĘZYKOWYCH C1 DOTYCZĄCYCH BIZNESOWEGO JĘZYKA ANGIELSKIEGO

Centrum Kształcenia Językowego Uniwersytetu Masaryka (CKJ UM) rozpoczęło w 2012 roku, zgodnie ze swoją koncepcją nauczania języków obcych, proces doskonalenia i standaryzacji sprawdzania i oceniania umiejętności językowych studentów kierunków nefilologicznych UM. Proces ten odbywa się na dwóch płaszczyznach. Pierwszą jest standaryzacja formy wyjściowych testów językowych na poziomach B1, B2 i C1 dla wszystkich języków nauczanych w CKJ UM. Drugą płaszczyzną standaryzacji polega na wdrożeniu dobrej praktyki rozwoju testu językowego w taki sposób. W kontekście tego procesu zaprezentowano udoskonalenie i standaryzację testów językowych sprawdzających biznesowy język angielski na poziomie C1. Wykorzystując wyniki wstępnego sprawdzenia w ten sposób zmodyfikowanych testów, przeprowadzona zostanie analiza ich wpływu na sprawdzaną populację pod kątem oczekiwanych korzyści. Zbadane zostaną także subiektywne oceny studentów dotyczące zmodyfikowanych testów pod kątem ich przydatności.