

Univerzita Karlova v Praze, Pedagogická fakulta



# **Zjišťování výsledků vzdělávání v anglickém jazyce**

**Disertační práce**

**PhDr. Mária Šikolová**

**Školitel: Doc. Ing. Petr Byčkovský, CSc.**

**Praha, 2007**

## Obsah

### SEZNAM VYBRANÝCH CÍLŮ A TEORETICKÉ KARLŮVY PŘEDVÁNÍ JAZYKŮ

Touto cestou bych ráda poděkovala svému školiteli Doc. Ing. Petru Byčkovskému, CSc. za jeho odborné vedení, osobní přístup a čas, který mi věnoval. Kromě toho děkuji svým kolegům a kolegyním za rady a připomínky, jakož i středoškolským učitelům, kteří mi umožnili ověření jazykových testů v praxi a také se k nim kriticky vyjádřili.

*Prohlašuji, že jsem tuto disertační práci vypracovala samostatně s použitím literatury, která je uvedena v seznamu.*

2007

*[Handwritten signature]*

## Obsah

I. FORMULACE VZDĚLÁVACÍCH CÍLŮ A TEORETICKÉ ZÁKLADY TESTOVÁNÍ JAZYKŮ.....	4
1. Úvod.....	4
2. Cíl výzkumu a výzkumné metody.....	6
3. Vzdělávací cíle a jejich role v hodnocení studijních výsledků .....	7
3.1 Některé teoretické pohledy na tvorbu vzdělávacích cílů .....	7
3.2 Co by měly splňovat katalogy požadavků a hodnotící nástroje.....	9
3.3 Standardy vzdělávání, katalogy požadavků a reforma maturitní zkoušky v ČR .....	17
4. Testování anglického jazyka v současnosti.....	20
5. Testování základních řečových dovedností.....	22
5.1 Testování receptivních dovedností .....	22
5.1.1 Definice receptivních dovedností .....	22
5.1.2 Definice testů receptivních dovedností .....	25
5.1.3 Charakteristika testování receptivních dovedností .....	26
5.1.4 Testovací techniky receptivních dovedností .....	30
5.2. Testování produktivních dovedností.....	38
5.2.1 Definice produktivních dovedností a jejich testů .....	38
5.2.2 Charakteristika testování produktivních dovedností.....	39
5.2.3 Testovací techniky produktivních dovedností .....	44
5.2.4 Hodnotící škály .....	50
6. Testové úlohy z pohledu Bloomovy taxonomie.....	52
II. NÁVRH ALTERNATIVNÍHO KATALOGU POŽADAVKŮ; KONSTRUKCE, ZADÁNÍ A VYHODNOCENÍ TESTU .....	58
7. Hodnocení Katalogu požadavků: pozitivní a negativní stránky.....	58
7.1 Seznámení se s problematikou – zjištění získaná v dosavadní práci .....	58
7.2 Jaké jsou pozitivní vlastnosti Katalogu požadavků? .....	58
7.3 Které charakteristiky Katalogu požadavků je potřeba zdokonalit? .....	60
7.3.1 Obecné formulace .....	60
7.3.2 Receptivní dovednosti.....	60
7.3.3 Produktivní dovednosti .....	61
7.3.4 Mediační dovednosti .....	64
7.3.5 Tématické okruhy .....	64
7.4 Jsou formulace v Katalogu požadavků jasné a srozumitelné?.....	65
7.5 Další otázky související s Katalogem požadavků.....	66

8. Návrh alternativního Katalogu požadavků pro státní maturitní zkoušku z anglického jazyka .....	68
8.1 Katalog požadavků pro žáky, rodiče a širokou veřejnost .....	68
8.1.1 Požadovaná úroveň řečových dovedností .....	68
8.1.2 Doporučení vhodných učebních materiálů .....	69
8.1.3 Požadavky na výkony v jednotlivých řečových dovednostech .....	70
8.1.4 Požadovaná znalost jazykových prostředků .....	79
8.1.5 Počet dílčích testů a časové limity .....	83
8.1.6 Testovací techniky .....	84
8.1.7 Typy textů .....	86
8.1.8 Instrukce k testovým úlohám a kritéria hodnocení .....	88
9. Hodnocení alternativního katalogu požadavků .....	89
10. Konstrukce testu podle alternativního katalogu požadavků .....	91
10.1 Poslech s porozuměním .....	92
10.2 Čtení s porozuměním .....	93
10.3 Psaní .....	93
10.4 Mluvení .....	94
11. Zadání a hodnocení subtestů .....	95
11.1 Zadání subtestů .....	95
11.2 Hodnocení testů .....	95
11.2.1 Subtest zjišťující úroveň čtení .....	95
11.2.2 Subtest hodnotící úroveň psaní .....	97
12. Interpretace výsledků subtestů .....	100
12.1 Subtest čtení s porozuměním .....	100
12.2 Subtest psaní .....	106
12.3 Subtest čtení a psaní .....	106
13. Závěr .....	107
III. PŘÍLOHY .....	111
Příloha P – 1 Pojmy používané v disertační práci a jejich definice .....	111
Příloha P – 2 Výsledky pozorování průběhu maturitní zkoušky .....	117
Příloha P – 3 Dotazník k hodnocení alternativního katalogu požadavků .....	121
Příloha P – 4 Subtest poslechu s porozuměním .....	125
Příloha P – 5 Subtest čtení s porozuměním .....	133
Příloha P – 6 Subtest psaní .....	142
Příloha P – 7 Subtest mluvení .....	145
Příloha P – 8 Hodnotící kritéria písemného projevu .....	147
Příloha P – 9 Hodnotící kritéria ústního projevu .....	149
Příloha P – 10 Průvodní instrukce pro zadávající a dotazník k subtestu čtení .....	152
Příloha P – 11 Průvodní instrukce pro zadávající a dotazník k subtestu psaní .....	154
Příloha P – 12 Počítačový výstup vyhodnocení testu .....	156

Literatura .....	162
Bibliografický záznam, anotace, klíčová slova .....	165
SUMMARY .....	166

## 1. Úvod

V současné době široké mezinárodní ekonomické, vědecké, vzdělávací a kulturní spolupráce se zvyšuje význam jazykové vzdělávání na školách i v rámci celoživotního vzdělávání. Zvláště cizích jazyků a zejména angličtiny se stále jedná o nejdůležitější předpoklad pro studium i práci doma i v zahraničí. V souvislosti s výukou cizích jazyků vynikají různé nové přístupy a školní, týkající se jak učení a vyučování, tak i hodnocení dosažené jazykové úrovni.

Kada Evropy se mobilizuje jazykově vzdělávat v Evropě a hodnocení její výsledky v září 1971, avšak až v roce 1978 byl vydán Společný evropský referenční rámec pro jazyky (CEFR). Tento dokument je velmi důležitým mezinárodním prahem, který umožňuje učitelům i studentům cizí jazyky vyučovat. Cílem tohoto dokumentu je nabídnout všestranný, transparentní a koherentní rámec pro jazykové vzdělávání a hodnocení. Byl zpracován při podpoře koordinace jazykové výuky, která je jedním z celků vyšší mobility, elektronické koordinace komunikace v rámci evropského rámce s celkovou dimenzí, aby byla přístupná informacím, motivující učení koordinovaně, široce pracovních úkolů a jejich vzájemné porozumění. V Evropské Unii Evropský rámec pro jazyky (CEFR) je 33. Tento dokument navrhuje využití evropského rámce pro jazykové vzdělávání a hodnocení, který je založen na celkové dimenzi (A) - (C) a doporučení mezinárodních úkolů, úkolů a úkolů.

V České republice stále jazykové vzdělávání podle celkové dimenzí dimenzí, která zahrnuje všechny jazykové úrovně, které jsou v současnosti v České republice. Tato úroveň je založena na celkové dimenzi (A) - (C) a doporučení mezinárodních úkolů, úkolů a úkolů. V současnosti je v České republice v oblasti jazykové výuky stále v platnosti evropský referenční rámec pro jazyky (CEFR) a jeho využití v České republice je stále v platnosti. V současnosti je v České republice v oblasti jazykové výuky stále v platnosti evropský referenční rámec pro jazyky (CEFR) a jeho využití v České republice je stále v platnosti. V současnosti je v České republice v oblasti jazykové výuky stále v platnosti evropský referenční rámec pro jazyky (CEFR) a jeho využití v České republice je stále v platnosti.

# 1. FORMULACE VZDĚLÁVACÍCH CÍLŮ A TEORETICKÉ ZÁKLADY TESTOVÁNÍ JAZYKŮ

## 1. Úvod

V současné době široké mezinárodní ekonomické, vědecké, vzdělávací a kulturní spolupráce se zvyšuje význam jazykového vzdělávání na školách i v rámci celoživotního vzdělávání. Znalost cizích jazyků a zejména angličtiny se stává jedním ze základních předpokladů pro studium i práci doma i v zahraničí. V souvislosti s výukou cizích jazyků vyvstává mnoho nových otázek a úkolů, týkajících se jak učení a vyučování, tak i hodnocení dosažené jazykové úrovně.

Rada Evropy se problémy jazykového vzdělávání v Evropě a hodnocením jeho výsledků zabývá již od roku 1971, avšak až v roce 2001 byl vydán Společný evropský referenční rámec pro jazyky (SEER). Tento dokument je velmi důležitým materiálem pro ty, kdo se cizím jazykům učí i pro ty, kteří cizí jazyky vyučují. Cílem tohoto dokumentu je nabídnout všeobecný, transparentní a koherentní rámec pro jazykové vzdělávání a hodnocení. Byl zpracován pro podporu intenzifikace jazykové výuky, která „je potřebná v zájmu vyšší mobility, efektivnější mezinárodní komunikace s ohledem na národní identitu a kulturní různorodost, pro lepší přístup k informacím, intenzivnější osobní komunikaci, zlepšení pracovních vztahů a hlubší vzájemné porozumění“ (Common European Framework for Languages, 2001, s. 5). Tento dokument nevnucuje jednotlivým zemím jednotné jazykové osnovy a zkoušky, spíše nabízí jasně formulovanou a podrobnou charakteristiku řečových dovedností na šesti úrovních (A1 – C2) a doporučuje různé komunikační úkoly, aktivity a strategie.

V České republice nemá jazykové testování vhodné pro současné podmínky dlouhou tradici. Zjišťování úrovně jazykových dovedností mimo základních, středních a vysokých škol poskytují státní a soukromé jazykové školy a dále pak zahraniční testovací instituce nabízející mezinárodně uznávané jazykové testy a certifikáty, např. TOEFL, FCE, PET, KET, CAE (podrobněji viz kap. 4). Přestože se jazyky vyučují na všech stupních vzdělávání, pro žádný z těchto stupňů zatím neexistují národní standardizované jazykové zkoušky. Standardizovaná jazyková zkouška chybí zejména na středních školách, kde by měla objektivně a spolehlivě posoudit úroveň řečových dovedností žáků na konci jejich středoškolského vzdělávání. Jinými slovy, je důležité, aby všichni žáci, kteří budou v takové zkoušce úspěšní, měli alespoň minimálně požadovanou úroveň jazykových dovedností. V současné době je totiž ještě stále maturitní zkouška z cizího jazyka zcela v kompetenci jednotlivých středních škol a tudíž požadavky ke zkoušce, jakož i zjišťování plnění těchto požadavků, nemůže být na stejné úrovni. Z toho plyne, že hodnocení maturitní zkoušky z jazyka určitou známkou nemá vysokou vypovídací hodnotu. K vyřešení této situace by měl napomoci projekt reformy maturitních zkoušek, který vznikl v Centru pro zjišťování výsledků vzdělávání (CERMAT) a který by měl být realizován již v roce 2007.

Povinná zkouška z cizího jazyka, která je součástí nové maturity, se bude skládat ze dvou částí - písemné a ústní. Písemná zkouška bude ověřovat úroveň čtení s porozuměním, poslechu s porozuměním, jazykové kompetence a psaní. Ústní zkouška se zaměří na ověřování úrovně dovednosti mluvení. Novou koncepci maturitní zkoušky z cizího jazyka lze považovat za vhodnou, i když není možné počítat s tím, že ihned po svém zavedení přispěje k převratným změnám ve výuce cizích jazyků nebo k výraznému zdokonalení výstupní jazykové úrovně žáků středních škol. Změna ve způsobu ověřování výstupních jazykových znalostí a dovedností je nepochybně změna správným směrem a měla by také mít pozitivní dopad na výuku cizích jazyků.

Tato práce byla započata v době, kdy byl vytvořen první *Katalog požadavků ke společné části maturitní zkoušky v roce 2004* (dále jenom Katalog požadavků), a proto se nezabývá novou koncepcí maturity. Její praktická část vychází z původního dokumentu. Cílem jeho vytvoření bylo jasně stanovit obsah a formu maturitní zkoušky, aby podmínky pro všechny žáky středních škol byly stejné a jejich výsledky srovnatelné. Předpokládalo se, že nová koncepce maturity bude zavedena již v roce 2004, ale z mnoha důvodů k tomu nedošlo. Zmíněný dokument rozlišoval dvě různé úrovně obtížnosti a stanovoval, že maturita bude mít tzv. společnou a školní část. Požadavky byly podrobně rozpracovány jenom pro část společnou (testování poslechu a čtení). Pro školní část (testování mluvení a psaní) byly formulovány pouze obecnější požadavky.

ČÁST I se zabývá vzdělávacími cíli, charakteristikou jednotlivých řečových dovedností a testů ověřujících jejich úroveň, jakož i testovacími technikami, které se v současnosti k jejich ověřování používají. Zmíněné testovací techniky používané v jazykovém testování se pak pokouší zařadit do Bloomovy taxonomie vzdělávacích cílů. Protože se práce zabývá zjišťováním výsledků vzdělávání, je stanovení požadavků ke zkoušce velmi důležité. Zkouška totiž pak ověřuje, zda předem stanovené požadavky byly či nebyly splněny. Proto jsou v první části uvedeny pohledy některých badatelů na danou problematiku. Pozornost je dále věnována základním vlastnostem, které by měl kvalitní test obecně mít (validita, reliabilita). Stručně je popsán současný stav reformy maturitní zkoušky v anglickém jazyce. Důraz je kladen na soudobé trendy v testování cizích jazyků. Protože se ve výuce jazyků v současnosti preferují komunikativní přístupy, odráží se tato skutečnost i v testování jazyků. To v praxi znamená, že se ustupuje od ověřování zejména nebo pouze znalostí „o jazyce“ a stále častěji se preferuje ověřování jednotlivých řečových dovedností (poslech, čtení, psaní, mluvení). Pojednání o jazykovém testování uvádí teoretické pohledy expertů na definování jednotlivých řečových dovedností, uvádí nejčastější typy testových úloh používaných v jazykovém testování a zamýšlí se nad jejich výhodami a nevýhodami. Oblasti, kterými se ČÁST I zabývá, představují relevantní teoretická východiska pro ČÁST II.

Cílem ČÁSTI II je zhodnotit již zmíněný dokument Katalog požadavků, který vydalo MSMT v roce 2000, a navrhnout alternativní formu tohoto katalogu (dále alternativní katalog), která by se pokusila vyřešit sporné stránky zmíněného dokumentu. Dalšími cíli této části je navrhnout test vycházející z alternativního katalogu, ověřit subtest zjišťující úroveň dovednosti čtení s porozuměním a psaní a získané výsledky popsat, analyzovat a interpretovat. Hodnocení Katalogu požadavků se zaměřuje na jeho pozitivní

charakteristiky, jakož i na charakteristiky, které by bylo vhodné zdokonalit. Zaujímá stanovisko k obecným formulacím a jejich srozumitelnosti a také k obsahu a metodám ověřování jednotlivých řečových dovedností. Návrh alternativního katalogu požadavků se snaží respektovat současné teoretické pohledy na testování cizích jazyků, specifika maturitní zkoušky a také uznávaný dokument popisující jednotlivé úrovně pokročilosti v řečových dovednostech, Společný evropský referenční rámec. Uvedený alternativní katalog pak slouží jako východisko pro návrh testu pro úroveň mírně pokročilou ověřující všechny základní řečové dovednosti. Po zadání a opravě subtestů ověřujících dovednosti čtení a psaní jsou výsledky analyzovány a interpretovány. Pro vyhodnocení testu čtení bylo použito softwarového programu ITEMAN.

## 2. Cíl výzkumu a výzkumné metody

### Cíl výzkumu

Jak bylo naznačeno v úvodu, cílem výzkumu je upravit Katalog požadavků a navrhnout jeho alternativní verzi, která by sloužila jako východisko k maturitní zkoušce z anglického jazyka a která by se zabývala formou a obsahem zkoušky.

### Výzkumné otázky

K uvedenému cíli se váže mnoho výzkumných otázek, z nichž podstatné uvádím:

1. Jak dosud probíhají maturitní zkoušky z anglického jazyka, co a jak se při nich zkouší?
2. Jaké jsou pozitivní charakteristiky Katalogu požadavků?
3. Které charakteristiky Katalogu požadavků je potřeba zdokonalit?
4. Co by měly zahrnovat jednotlivé řečové dovednosti?
5. Jaké by měly být minimální výkony v jednotlivých řečových dovednostech?
6. Co a na jaké úrovni by měli studenti znát z jazykových prostředků?
7. Jak ověřit subtesty, jak je vyhodnotit a jak interpretovat výsledky?

### Výzkumné metody

Při hledání odpovědí na výše uvedené výzkumné otázky bylo použito různých výzkumných metod.

1. Jak dosud probíhají maturitní zkoušky z anglického jazyka, co a jak se při nich zkouší?

*Výzkumné metody:*

kvalitativní analýza:

- pozorování průběhu maturitní zkoušky
- zčásti strukturované rozhovory s učiteli středních škol

2. Jaké jsou pozitivní charakteristiky Katalogu požadavků?
3. Které charakteristiky Katalogu požadavků je potřeba zdokonalit?



*Výzkumná metoda:*

- obsahová analýza

4. Co by měly zahrnovat jednotlivé řečové dovednosti?
5. Jaké by měly být minimální výkony v jednotlivých řečových dovednostech?
6. Co a na jaké úrovni by měli studenti znát z jazykových prostředků?

*Výzkumná metoda:*

- vývoj alternativního katalogu na základě aplikovaného výzkumu

7. Jak ověřit subtesty a jak je vyhodnotit a interpretovat výsledky?

*Výzkumné metody:*

kvantitativní analýza:

- ověření v terénu
- ověření empirických dat
- položková analýza

### **3. Vzdělávací cíle a jejich role při hodnocení výsledků vzdělávání**

Aby bylo dosaženo cíle výzkumu, je potřeba seznámit se se stavem bádání v relevantní teoretické oblasti. Tato kapitola pojednává o vzdělávacích cílech, o jejich formulaci a o způsobech, jak je využívat pro tvorbu testů. Zmiňuje také nejdůležitější skutečnosti týkající se této problematiky v českém středním školství. Definice pojmů používaných v práci i s anglickými ekvivalenty jsou uvedeny v Příloze P - 1.

#### **3.1 Některé teoretické pohledy na formulování a klasifikaci vzdělávacích cílů**

Důležitost vzdělávacích cílů určitě nepochybně nikdo, kdo pracuje v oblasti vzdělávání. Otázkou však zůstává, jak cíle správně formulovat, jak umožnit ve výuce jejich realizaci a, v neposlední řadě, jak míru jejich plnění přesně a spolehlivě změřit.

Problematikou vzdělávacích cílů se zabývalo i zabývá mnoho odborníků z oblasti pedagogiky i psychologie. Vzhledem k rozsahu a zaměření práce si stručně všimneme jen prací tří zástupců pedagogů, kteří se problematikou cílů zabývali a jejichž práce patří ve světě k nejznámějším. Jedná se o Roberta F. Magera, Benjamina S. Blooma a jeho spolupracovníky a následovníky a Jonathana Muellera.

R.F. Mager formuluje tři hlavní důvody, proč se vzdělávacími cíli zabývat. Jako první uvádí skutečnost, že pokud neexistují správně definované cíle, je velmi problematické vybírat výukové materiály, rozhodovat se o obsahu nebo určovat vhodné metody vyučování. Dalším důvodem, proč vzdělávacím cílům věnovat pozornost, je potřeba zjišťování, zda cíle bylo dosaženo. Pokud tedy neexistuje jasný cíl, testy jsou „přinejmenším zavádějící“. Za třetí nejdůležitější důvod k formulaci vzdělávacích cílů pak Mager považuje to, že poskytují studentům „prostředek k uspořádání vlastní snahy o dosažení těchto cílů“ (Mager, 1975, s. 5-6). Všechny uvedené důvody jsou nepochybně velmi důležité, pro tuto práci však za nejvíce relevantní lze pokládat druhý důvod –

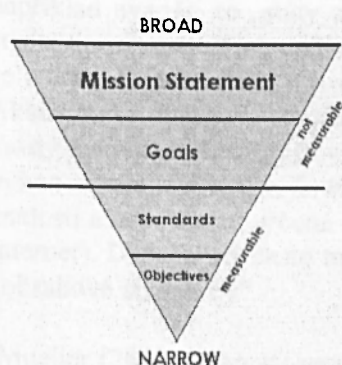
potřeba zjišťovat, zda bylo stanoveného cíle dosaženo. Ve svých úvahách o vzdělávacích cílech jde Mager dále do hloubky a dokonce uvádí seznam činných sloves, která se často ve formulacích cílů vyskytují a rozděluje je na slovesa, jejichž význam dává prostor k mnoha interpretacím a slovesa, jejichž význam umožňuje méně interpretací. Právě druhou skupinu pokládá za vhodnější při stanovování vzdělávacích cílů (Mager, 1975, s. 20). Nosnou myšlenkou jeho publikace *Preparing Instructional Objectives* je však charakteristika užitečného vzdělávacího cíle, který by podle něj měl obsahovat:

- „1. Výkon (*Performance*). Cíl by měl vždy říci, co má být schopen student udělat.
2. Podmínky (*Conditions*). Cíl vždy popisuje důležité podmínky (pokud nějaké existují), za kterých má být výkon proveden.
3. Kritérium (*Criterion*). Kdykoliv je to možné, cíl uvádí kritérium přijatelného výkonu. Toto kritérium popisuje, jak dobře musí učící se provést výkon, aby byl přijatelný“ (Mager, 1975, s. 21).

Problematikou vzdělávacích cílů se také zabýval B.J. Bloom a tým jeho spolupracovníků, kteří společně vytvořili taxonomii vzdělávacích cílů. Taxonomii vzdělávacích cílů v kognitivní oblasti lze definovat jako hierarchicky uspořádaný rámec, který umožňuje klasifikovat cíle podle jejich kognitivní úrovně. Publikace Blooma a jeho týmu zabývající se touto problematikou byla vydána již v roce 1956. Prvním stimulem k jejímu vypracování bylo setkání středoškolských a vysokoškolských examinátorů v Bostonu již v roce 1948. Protože velká část hodnocení se v USA prováděla a provádí pomocí testových položek s výběrem odpovědi z vícenásobné volby a příprava tohoto typu položek je velmi náročná, záměrem pracovní skupiny bylo vzájemně si při jejich tvorbě pomáhat. Za tímto účelem bylo potřeba vytvořit určitý společný rámec, do něhož by bylo možné každou položku zařadit, a to podle toho, co měří. Takový byl tedy prvotní záměr, ze kterého časem vzešla v současnosti obecně známá Bloomova taxonomie. Od svého vzniku se stala velmi uznávanou a byla přeložena do více než dvaceti jazyků. V roce 2001 vychází nová, rozšířená verze, která vznikla pod vedením jednoho z týmu autorů původní verze D.R. Krathwohla.

Nová verze taxonomie se snaží pomoci při hledání odpovědí na čtyři základní otázky:

- „1. Co je důležité, aby se studenti naučili ve škole a ve třídě v čase, který je k dispozici a je omezený? (otázka učení – *learning question*)
2. Jak se má plánovat a provádět výuka, aby vedla k efektivnímu učení velkých počtů studentů (otázka vyučování – *instruction question*)
3. Jak se mají vybírat nebo navrhnout nástroje hodnocení a postupy, které poskytnou přesnou informaci o tom, jak dobře se studenti učí? (otázka hodnocení – *assessment question*)
4. Jak je možné zajistit, aby cíle, výuka a hodnocení byly vzájemně konzistentní? (otázka propojení – *alignment question*),“ (Anderson; Krathwohl 2001, s. 6).



Za zajímavou lze v této souvislosti pokládat i představu Jonathana Muellera o vzdělávacích standardech a cílech. Za nejširší a poměrně stručnou formulaci pokládá programové prohlášení určité vzdělávací instituce nebo státního vzdělávacího systému (*mission statement*), které velmi obecně uvádí, co „mají studenti znát a umět, když ukončují vzdělání“. Příkladem ve vyučování anglického jazyka by mohlo být například: „studenti umí komunikovat v anglickém jazyce.“ Vzdělávací cíle (*goals*) jsou pak již sice konkrétnější, ale jsou stále ještě vyjádřeny dost široce, například „studenti umí ústně i písemně komunikovat v běžných situacích každodenního života.“

Obr. 1. Grafické znázornění vzdělávacích dokumentů popisujících cíle vzdělávání (Mueller, 2006, internet).

Standardy (*standards*) se v uvedeném grafickém znázornění nacházejí mezi vzdělávacími cíli (*goals*) a výukovými cíli (*objectives*), což ukazuje, že standardy jsou formulovány obecněji než výukové cíle, ale konkrétněji než vzdělávací cíle. Jako příklad *standardu* je možné uvést: „student umí identifikovat hlavní myšlenky textu“. Za nejkonkrétnější můžeme označit výukový cíl, který lze formulovat například takto: „student umí vytvořit jednoduché otázky.“

Důležitou charakteristikou programového prohlášení a vzdělávacích cílů je to, že nejsou měřitelné, zatímco standardy a výukové cíle se označují jako měřitelné (Mueller, 2006, internet).

V souladu s popisem rozdílů mezi vzdělávacími cíli (*goals*) a výukovými cíli (*objectives*) Muellera jsou i další definice uveřejněné na internetu – mnohé z nich je možné nalézt na webových stránkách různých vzdělávacích institucí (<http://www.personal.psu.edu/staff/b/x/bxb11/Objectives/index.htm>; <http://www.pitt.edu/~ciddeweb/FACULTY-DEVELOPMENT/FDS/goalsobi.html>; [http://www.lib.utexas.edu/services/instruction/tips/cp/cp\\_goals.html](http://www.lib.utexas.edu/services/instruction/tips/cp/cp_goals.html)).

Z uvedeného vyplývá velká důležitost formulace jasných vzdělávacích cílů obecně, což platí samozřejmě i pro oblast výuky jazyků. Orientace v této oblasti je nezbytná pro splnění cíle výzkumu (kap. 2). Další subkapitoly se budou zabývat tím, jak se uvedené teoretické myšlenky mají uplatňovat v pedagogické praxi.

### 3.2 Co by měly splňovat katalogy cílových požadavků a hodnotící nástroje

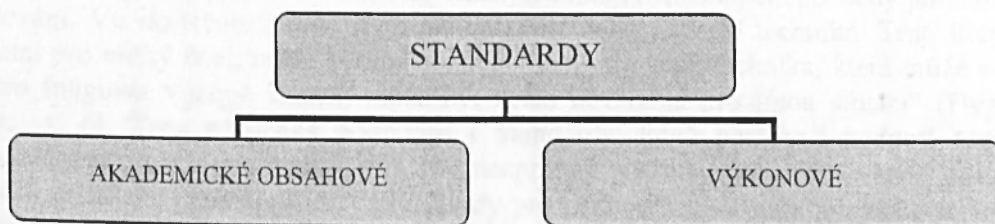
V předešlé subkapitole byly zmíněny vzdělávací cíle, jejich role i doporučení, co by měly obsahovat a jak je správně formulovat. Kromě toho byl uveden grafický model, který ilustruje postavení standardů v hierarchii vzdělávacích dokumentů. Vzdělávacím standardům se v současné době věnuje velká pozornost, a to jak v zahraničí, tak i v České

republice. Webové stránky zabývající se státním vzděláváním ve Spojených státech například uvádějí že „státy začaly používat akademické standardy, aby objasnily, co se studenti mají naučit a co učitelé mají vyučovat“ (1996, internet). Dále je možné seznámit se s definicí akademických obsahových standardů a s definicí výkonových standardů. Akademické obsahové standardy (*academic content standards*) specifikují to, co „by měl každý student znát a umět v základních akademických obsahových oblastech“, zatímco výkonové standardy (*performance standards*) popisují, jak dobře mají studenti zvládnout znalosti a dovednosti určené obsahovými standardy (Improving America's School, 1996, internet). Dnes se již často místo „akademické obsahové standardy“ používá zkráceného „obsahové standardy“.

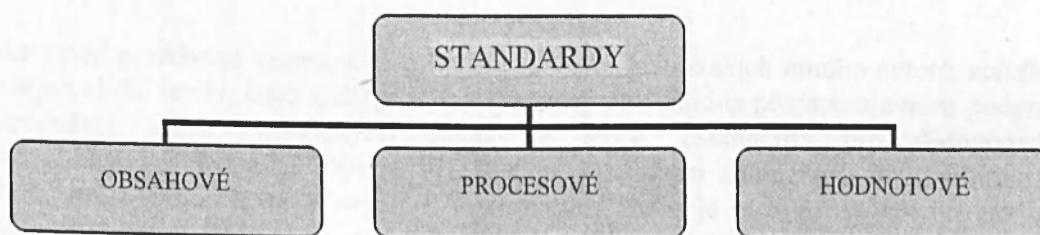
Mueller (2006, internet) ve shodě s uvedenými definicemi rozlišuje standardy obsahové (*content standards*) a jejich vymezení chápe stejně jako předchozí uvedení autoři; jako další typy standardů však uvádí standardy procesové (*process standards*) a standardy hodnotové (*value standards*).

Své rozdělení standardů zdůvodňuje tím, že samotné obsahové standardy nemohou pokrýt např. „dovednosti kritického myšlení, schopnosti řešit problém (procesuální standardy) nebo dovednost spolupracovat (hodnotový standard)“. Je však diskutabilní, zda Muellerem uvedené příklady procesuálních a hodnotových standardů nepatří mezi vzdělávací cíle.

Srovnání rozdělení standardů podle dokumentu *Improving America's School* a podle Muellera ilustrují následující obrázky.



Obr. 2. Rozdělení standardů podle dokumentu *Improving America's School* (1996, internet).



Obr. 3. Rozdělení standardů podle Muellera (Mueller, 2006, internet).

Jak tvorba standardů, tak navržení odpovídajících testů, jejich administrace, vyhodnocení i interpretace výsledků by se měly řídit určitými zásadami. K provádění tohoto náročného procesu profesionálním způsobem nedávno vyšla ve Spojených státech publikace *Standards for Educational and Psychological Testing* (Standardy pro pedagogické a psychologické testování, dále Standardy), která byla přeložena také do češtiny. V jejím chápání nabývá slovo standardy ještě dalšího významu – představuje principy, které by v procesu testování měly být dodržovány. Hned v úvodu se uvádí, že „účelem vydání *Standardů* je vymezit kritéria pro hodnocení testů, postupů při testování a výsledného dopadu užívání testů“ (Standardy pro pedagogické a psychologické testování 2001, s. 10). Mnoho jiných publikací zabývajících se testy a testováním některé ze zásad, které je nutné v procesu testování uplatňovat, zmiňuje, nicméně uvedená publikace se zabývá výhradně touto problematikou a postihuje ji do hloubky.

Jakými úvahami tedy při tvorbě standardizovaného testu začít a jakými zásadami se řídit?

Vzhledem k tomu, že proces testování od vývoje testu přes administraci k jeho vyhodnocení a vyvození závěrů je velmi komplikovaný, pracovníci v testování musejí mít neustále na zřeteli, že neexistuje žádný ideální test. Na druhé straně však „existují rozsáhlé důkazy svědčící pro efektivitu dobře vytvořených testů použitých ve vhodném kontextu“ (Standardy pro pedagogické a psychologické testování 2001, s. 8). Kromě testu samotného je důležitá také situace, ve které se test použije. Test navržený pro určitou populaci za určitým účelem není většinou automaticky vhodný pro jinou populaci, která je testována za jiným účelem. Tuto skutečnost vyjadřuje velmi pregnantně Hughes: „Někdy jsou examinátoři jazyků dotazováni, jaký je „nejlepší test“ nebo „nejlepší testovací technika“. Takovéto otázky odhalují nepochopení podstaty jazykového testování. Ve skutečnosti neexistuje nejlepší test nebo nejlepší technika. Test, který je ideální pro určitý účel, může být nepoužitelný pro jiný účel; technika, která může velmi dobře fungovat v jedné situaci, může být zcela nevhodná pro jinou situaci“ (Hughes, 1992, s. 6). Tuto myšlenku podporují i Standardy, když pozitivně hodnotí správné používání testů, ale zároveň varují, že „nesprávné použití testů může vážně poškodit osobu, která test podstoupila...“ (Standardy pro pedagogické a psychologické testování 2001, s. 8).

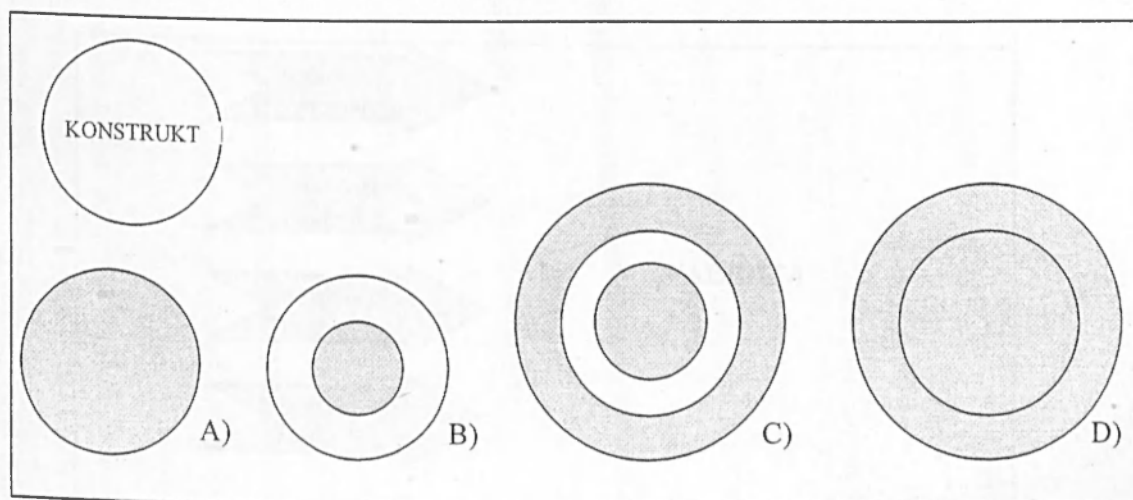
Nyní budou zmíněny některé důležité charakteristiky, jimiž by standardizovaný test vysoké důležitosti měl disponovat. Za základní vlastnosti dobrého testu lze označit validitu a reliabilitu; kromě toho bude zmíněno i několik dalších doporučených charakteristik.

Jako první potřebnou vlastnost testu uvádí ve svých publikacích mnoho autorů *validitu*. Test je validní tehdy, když zjišťuje to, co zjišťovat má. Validita představuje míru podpory interpretací skóre empirickými důkazy a teorií. (Standardy pro pedagogické a psychologické testování 2001, s. 17). Bachman a Palmer zdůrazňují, že „...validita se vztahuje k smysluplnosti a vhodnosti *interpretací*, které je možné odvodit na základě výsledků testu“ (Bachman; Palmer 1996, s. 21). Pokud tedy určitý počet bodů u jazykového testu interpretujeme určitým způsobem, musíme mít také argumenty, kterými můžeme naše tvrzení doložit. S validitou úzce souvisí definice *konstruktů*.

Konstrukt vyjadřuje to, co testem chceme měřit. Test by pak měl být navržen v souladu s proklamovaným konstruktem. Na základě konstruktů a výsledků testu je pak možné usuzovat, co je schopen kandidát v praxi vykonávat. V souvislosti s validitou je kromě jasné definice konstruktů důležité také stanovit, *k čemu budou výsledky testu použity* a jak budou interpretovány - jestli výsledky testu rozhodnou o zařazení do kurzu jisté pokročilosti nebo o splnění či nesplnění požadavků k přijetí na vysokou školu či do zaměstnání.

Není jednoduché zkonstruovat test, který by plně odpovídal proklamovanému konstrukt. Dva základní problémy, které mohou vzniknout, jsou definovány jako nedostatečná reprezentace konstruktů (*construct underrepresentation*) a rozptyl nevztahující se ke konstruktům (*construct irrelevance*). Nedostatečná reprezentace konstruktů znamená, že testové položky nepokrývají všechny důležité části konstruktů a význam testových skóre se tudíž zužuje (Standardy pro pedagogické a psychologické testování 2001, s. 18). V oblasti hodnocení jazykové úrovně si lze takovýto nedostatek představit například jako zařazení nedostatečně rozmanitých poslechových textů, které nepokrývají všechny tematické oblasti, nebo zařazení textů, které neodpovídají charakteristice kandidátů (věk, vzdělání) apod. Rozptyl nevztahující se ke konstruktům je možné definovat jako „míru, ve které jsou testové skóre ovlivněny procesy nesouvisejícími s předpokládaným konstruktem“ (Standardy pro pedagogické a psychologické testování 2001, s. 18). Častou chybou jazykových testů je použití testových úloh, které kromě úrovně řečových dovedností testují také paměť, případně faktické znalosti, aniž by byly uvedeny ve specifikacích a aniž by si tvůrci testů uvědomili, že je testují.

Následující obrázek znázorňuje konstrukt a dva základní problémy, které se mohou při tvorbě testu vyskytnout – nedostatečná reprezentace konstruktů a rozptyl nevztahující se ke konstruktům.



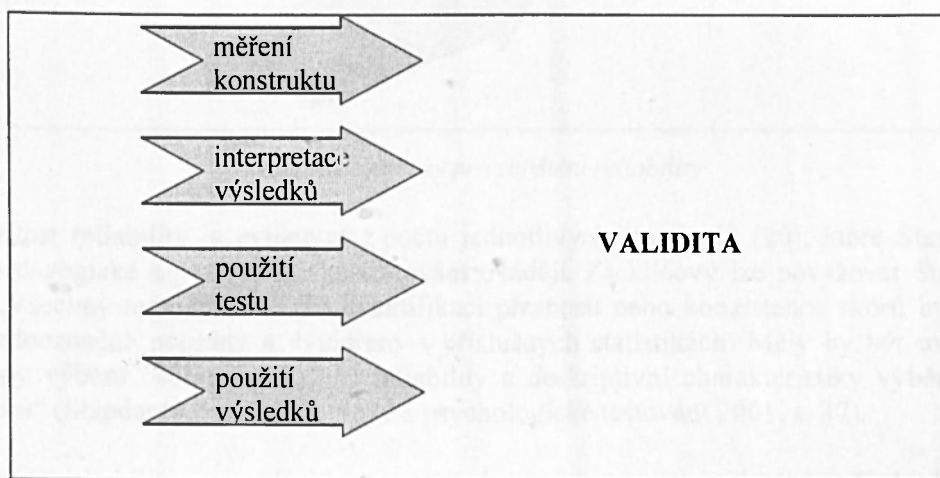
Obr. 4. Reprezentace konstruktů v testu

*Legenda k obr. 4:*

- A - ideální test, reprezentující celý konstrukt,
- B - test nedostatečně (neúplně) reprezentující konstrukt,
- C - test nedostatečně (neúplně) reprezentující konstrukt a navíc obsahující rozptyl nevztahující se ke konstrukt,
- D - test reprezentující celý konstrukt, ale obsahující i rozptyl nevztahující se ke konstrukt.

K tomu, aby tvůrci testu mohli tvrdit, že jimi navržený test je validní, je potřeba uvedení důkazů o validitě. Existuje mnoho zdrojů důkazů, které mohou zkoumat validitu testu. Budou zmíněny pouze dva, které lze považovat za klíčové. První z nich se opírá o zjišťování vztahu mezi obsahem testu a konstruktem. V ideálním případě by konstrukt měl být plně reflektován v obsahu testu a také by mu měly odpovídat typy testových úloh. Další se týká důkazů založených na postupech odpovídání na testové položky. Kandidáti jsou dotazováni na strategie řešení testových úloh, které používali. Tato oblast se netýká pouze kandidátů, ale také hodnotitelů – „pokud mají ...posuzovatelé používat při hodnocení výkonu testované osoby určitá kritéria, je důležité se ujistit, že tato kritéria skutečně uplatňují a nejsou ovlivňováni faktory nesouvisejícími se zamýšlenou interpretací“ (Standardy pro pedagogické a psychologické testování 2001, s. 20). Správné používání hodnotících kritérií je zejména pro oblast jazykového testování velmi důležité, protože hodnocení mluvení a psaní obvykle používá subjektivní metody a proto je potřeba usilovat o co největší konzistentnost hodnocení různých hodnotitelů.

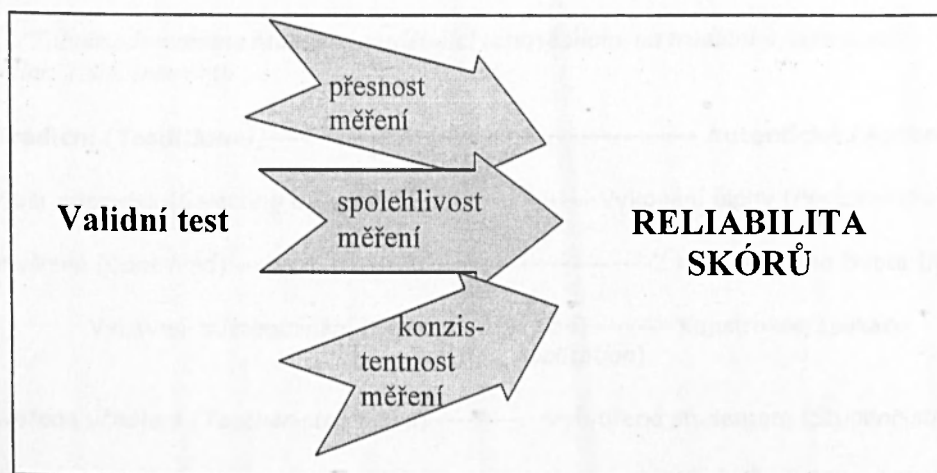
Následující obrázek shrnuje předchozí část týkající se validity a znázorňuje nejdůležitější podmínky, jejichž splněním je validita podmíněna. Pro dosažení validity je tedy potřeba, aby test měřil, co měřit má (to, co je uvedeno v konstrukt), aby výsledky testu byly správně interpretovány, aby byl test použitý pro účel, pro který byl zkonstruován, a aby se výsledky testu správně použily.



*Obr. 5. Podmínky pro zajištění validity*

Další z vlastností, kterými by dobrý test měl disponovat, je *reliabilita (spolehlivost)*. V praxi to znamená, že výsledky testu jsou přesné a spolehlivé: kdyby se stejným studentům zadal stejný test dvakrát za stejných podmínek, v ideální situaci by výsledky testu měly být totožné. Jinými slovy je reliabilitu možné definovat jako konzistentnost měření. Chráska uvádí, že reliabilita testu závisí na kvalitě i kvantitě testových položek. Vyšší počet testových položek obecně zvyšuje reliabilitu. Chráska dále upozorňuje, že test musí být dostatečně validní, aby disponoval vysokou reliabilitou, ale „vysoká reliabilita didaktického testu však ještě není zárukou toho, že test bude validní.“ (Chráska, 1999, s. 18). Také Standardy zdůrazňují potřebu přesnosti a konzistentnosti měření. „Přesnost je ale tím důležitější, čím závažnější důsledky vyplývají z interpretace naměřených skóre a čím významnější rozhodnutí je na nich založeno“ (Standardy pro pedagogické a psychologické testování 2001, s. 35). S tímto tvrzením souhlasí i Chráska, když uvádí, že nízká reliabilita testu vede ke skeptičtějšímu posuzování výsledků testu (Chráska, 1999, s. 18).

Následující obrázek schematicky uvádí, jaké podmínky musí být splněny pro zajištění reliability testových skóre. První podmínkou je validita testu, poté je potřeba, aby měření byla přesná, spolehlivá a konzistentní.



Obr. 6. Podmínky pro zajištění reliability

Důležitost reliability je evidentní z počtu jednotlivých standardů (20), které Standardy pro pedagogické a psychologické testování uvádějí. Za klíčový lze považovat Standard 2.4: „Všechny metody užití při kvantifikaci přesnosti nebo konzistence skóre by měly být jednoznačně popsány a vyjádřeny v příslušných statistikách. Měly by být uvedeny postupy výběru osob pro analýzu reliability a deskriptivní charakteristiky výběrového souboru“ (Standardy pro pedagogické a psychologické testování 2001, s. 37).

Problém reliability v jazykovém testování může nastat zejména při hodnocení produktivních dovedností mluvení a psaní. Pro zajištění vysoké reliability je nutné zpracovat a v praxi ověřit jasná a jednoznačná hodnotící kritéria, která umožní různým jednotlivcům hodnotit podle stejného měřítka. Kromě toho je důležité hodnotitele



odpovídajícím způsobem vyškolit a poté neustále monitorovat jak konzistentnost hodnocení různých jednotlivců mezi sebou navzájem (*inter-rater reliability*), tak i konzistentnost hodnocení jednotlivců (*intra-rater reliability*).

Kromě reliability a validity si u testu rovněž všímáme *objektivitu*. Objektivita testu záleží na testovacích technikách; například při použití techniky testování s výběrem odpovědi z vícenásobné nabídky a, b, c, d je test hodnocen objektivně. Pokud se používají metody testování s otevřenými odpověďmi, je potřeba vytvořit podmínky k tomu, aby byl test hodnocen objektivně – tedy různými osobami stejně; u jazykového testování je objektivita hodnocení neproblematická například u testových úloh s výběrem odpovědi nebo u dichotomických úloh, které se většinou používají pro hodnocení receptivních dovedností. Při hodnocení produktivních řečových dovedností psaní a mluvení může být zajištění objektivitu hodnocení problematictější.

Další otázkou zkoumanou v souvislosti s testováním je *autenticita*. Mueller charakterizuje autentickou úlohu jako „úkol zadaný studentům určený k hodnocení jejich schopnosti použít znalosti a dovednosti formulované ve standardech k řešení každodenních problémů“. Zajímavé je jeho rozdělení testových úloh na ose dvou protikladů - *tradiční* a *autentické*, které ilustruje tab.1.

Tab.1. Tabulka Jonathana Muellera rozdělující testové úlohy na tradiční a autentické. (Mueller, 2006, internet).

<b>Tradiční (Traditional)</b>	<b>Autentické (Authentic)</b>
Výběr odpovědi ( <i>Selecting a Response</i> )	Vykonání úlohy ( <i>Performing a Task</i> )
Vymyšlené ( <i>Contrived</i> )	Z každodenního života ( <i>Real-life</i> )
Vybavení si/rozpoznání ( <i>Recall/Recognition</i> )	Konstrukce/aplikace ( <i>Construction/Application</i> )
Vytvořené učitelem ( <i>Teacher-structured</i> )	Vytvořené studentem ( <i>Student-structured</i> )
Nepřímý důkaz ( <i>Indirect Evidence</i> )	Přímý důkaz ( <i>Direct Evidence</i> )

Toto rozdělení je velmi zajímavé i z hlediska testování řečových dovedností. Bylo by nepochybně zajímavé zkoumat vliv, jaký může mít použití autentických testových úloh na jednotlivé charakteristiky testu, nicméně v jazykovém testování se v současnosti používají ještě stále i tradiční testové úlohy.

Mezi odborníky na jazykové testování se často diskutuje, i když méně píše, o důležitosti autenticity. Na problematiku autenticity je možné nahlížet z různých úhlů. V principu se však v jazykovém testování jedná zejména o autenticitu jazykového materiálu použitého v testu a o autenticitu testové úlohy. Autenticita souvisí s validitou testu. Například v jazykovém testu pro číšníky nebo recepční by nebylo napsání eseje na téma *Trest smrti ano nebo ne* otázkou autentickou. Podobně pro testování stejných profesních skupin by

doplňování klíčových frází do dialogu se zákazníkem bylo autentické co do situace nebo kontextu (rozhovor se zákazníkem, restaurace/recepce), nebylo by však autentické z hlediska typu testové úlohy (nepředpokládá se, že by číšníci nebo recepční v reálném životě takovouto činnost vykonávali). Autenticita představuje „korespondenci testových úloh a úloh v použití cílového jazyka v praxi“ (Bachman; Palmer 1996, s. 23-25).

Někteří autoři se při úvahách o testování také zmiňují o *dopadu testů (impact)* „na jednotlivce, vzdělávací systémy a na společnost obecně“ (Davies et al., 1999, s. 79). Jedním z možných dopadů testování je jeho vliv na vyučování. *Dictionary of Language Testing* definuje tento jev jako: „washback/backwash: vliv (efekt) testování na vyučování. Zpětný vliv jazykového testu je buď pozitivní nebo negativní“ (Davies et al., 1999, s. 225). Jako příklad negativního zpětného efektu pak slovník uvádí situaci, kdy se dovednost psaní zjišťuje pouze testovací technikou výběru z několikanásobné volby. To pak často vede učitele k tomu, aby se při vyučování spíše než na rozvoj dovednosti psaní zaměřili na procvičování typů položek, které se vyskytují v testu. Naopak příkladem pozitivního zpětného efektu při výuce jazyků může být zavedení ústního rozhovoru do závěrečné zkoušky, protože může podpořit procvičování konverzace ve výuce jazyka. V definici termínu je však uvedeno i upozornění, že pro mnohá tvrzení týkající se pozitivního i negativního zpětného efektu existuje pouze málo empirických důkazů (Davies et al., 1999, s. 225). Důležitost zpětného efektu uvádí i Davidson a Lynch, kteří se domnívají, že test není pouze testem sám o sobě, ale představuje také dopad na lidi, kteří ho používají (Davidson; Lynch 2002, s. 1).

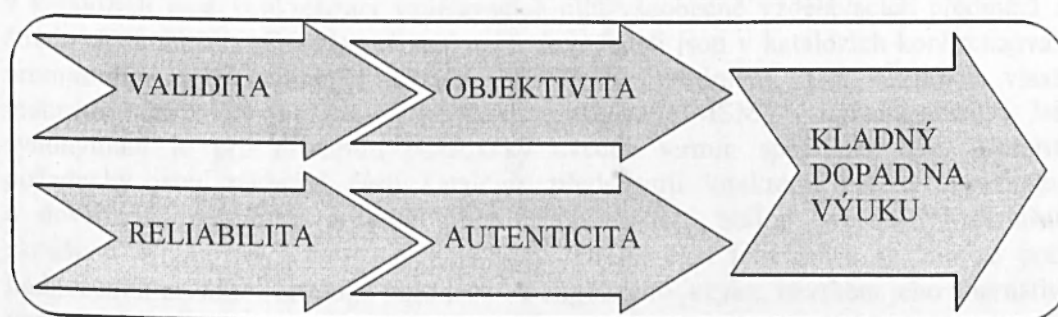
Obecně se pokládá za žádoucí, aby test měl na vedení výuky kladný dopad. Běžně se zkoumá dopad testů jak na jednotlivce, tak i na vzdělávací systémy. Jde teda vlastně o vliv testů na učení a vyučování. Hughes upozorňuje na to, že se při testování často uchylujeme k testování toho, co je nejjednodušší testovat, spíše než bychom se zabývali testováním toho, co je důležité testovat. To, co se netestuje, bude pravděpodobně v učení a vyučování zanedbáváno. Proto je důležité napsat testové specifikace přesně a v testu je plně aplikovat (Hughes, 1992, s. 23). Nad použitím testů jako nástroje ovlivnění výuky žadaným směrem se zamýšlí také *Shohamy*. Jako příklad použití testu k ovlivnění výuky uvádí zavedení ústní zkoušky z anglického jazyka na izraelských středních školách. Ministerští úředníci chtěli, aby zavedení této zkoušky přimělo učitele angličtiny zvýšit důraz na vyučování mluveného jazyka, a tedy zvýšit úroveň ústního projevu u maturantů. Zkouška se tak stala metodou, jak získat kontrolu nad vyučováním (Shohamy, 2001, s. 75). Shohamy se také zmiňuje o některých hypotézách týkajících se zpětného efektu, které prezentovali Alderson a Wall. Jedna z jejich hypotéz uvádí, že „test může ovlivnit vyučování i učení: co učitelé vyučují, jak učitelé vyučují, co se žáci učí, jak se žáci učí, tempo a posloupnost učení, postoj k obsahu a metodám vyučování a učení“ (Alderson; Wall v Shohamy 2001, s. 48).

V souvislosti s problematikou vlivu testování na vyučování uvádějí Standardy, že někdy je možné prosazovat některé školní testy proto, aby se zlepšila motivace studentů či změnily vyučovací metody (Standardy pro pedagogické a psychologické testování 2001, s. 24).

Obecně pro testování všech řečových dovedností také platí, že je vždy lepší používat pro testování jedné řečové dovednosti více **různých testových technik**. Přestože například technika výběru správné odpovědi z vícenásobné nabídky nebo technika s dichotomní odpovědí hodnotí kandidáty úplně objektivně, neznamená to, že celý test by měl obsahovat pouze tuto/tyto techniku/y. Takovéto rozhodnutí by pak mohlo vést k negativnímu dopadu na výuku v tom smyslu, že by kandidáti nebyli vedeni k procvičování určité řečové dovednosti, ale spíše k procvičování této testovací techniky. „Obecně lze říci, že čím více různých metod test využívá, tím více můžeme věřit, že test není zaujatý vůči určité metodě nebo určitému druhu kandidáta“ (Alderson; Clapham; Wall 1995, s. 45).

Z uvedeného je zřejmé, že tvorba standardů a proces tvorby, zadávání a vyhodnocování testů vhodným způsobem jsou velmi náročné a vyžadují tým vyškolených pracovníků. Jenom tak může testování poskytnout správné a vhodně interpretované informace. Vzdělávací cíle a standardy by tedy neměly být pouze formálním, neživým dokumentem, ale měly by pomáhat ve výběru učiva, výukových metod i v konstrukci hodnotícího nástroje. Kromě toho je žádoucí, aby sloužily k jasnější orientaci v rámci jednotlivých předmětů jak učitelům, tak učícím se. Protože katalogy požadavků k maturitní zkoušce tyto cíle formulují a zároveň uvádějí i charakteristiku testů, jsou pro učitele, žáky i širokou veřejnost velmi důležitými a přínosnými dokumenty.

Základní charakteristiky dobrého standardizovaného testu ilustruje následující obrázek.



Obr. 7. Důležité charakteristiky dobrého standardizovaného testu

### 3.3 Standardy vzdělávání, katalogy požadavků a reforma maturitní zkoušky v ČR

Pokud se vrátíme k obr. 1 v subkapitole 3.1, je možné hledat paralely vzdělávacích dokumentů v České republice, které odpovídají uvedenému schématu. Za nejobecnější dokument formulující nejšířší cíle vzdělávání a strategie rozvoje školství lze pokládat Národní program rozvoje vzdělávání v České republice (Bílá kniha), který byl vydán v roce 2001. Vzhledem k tomu, že tato práce se koncentruje na maturitní zkoušku, tedy na střední školství, dalším důležitým dokumentem je Standard vzdělávání ve čtyřletém gymnáziu, který byl publikován již v roce 1996 ve Věstníku MŠMT, a který obsahuje

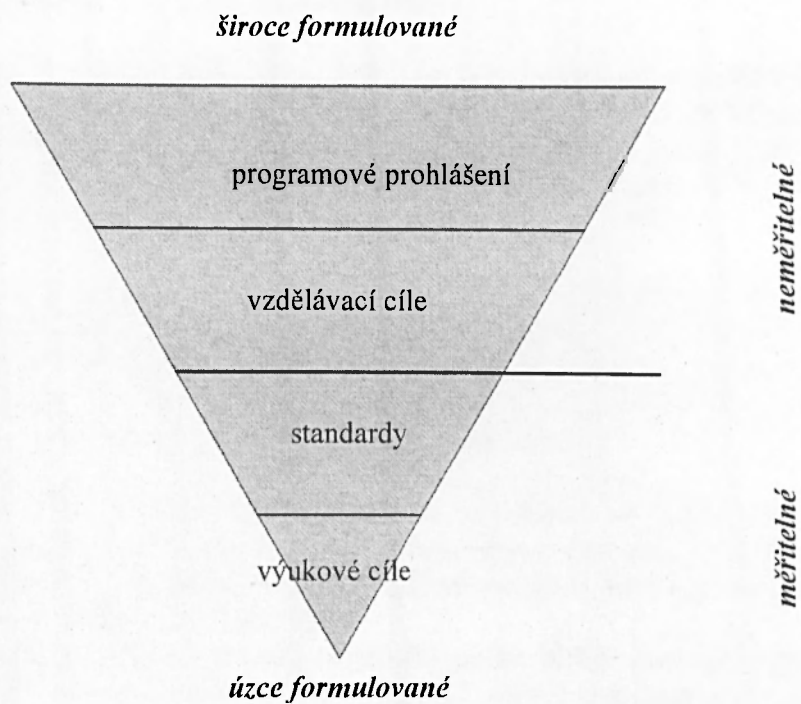
dvě hlavní části. První část formuluje pro absolventy čtyřletého gymnázia požadovaný rozvoj poznání, dovednosti a kompetence a hodnoty a postoje. Druhá část se zaměřuje na obsah kmenového učiva. Pro obor cizích jazyků pak formuluje specifické cíle a okruhy kmenového učiva. Tento dokument lze určitě označit za konkrétnější než Národní program rozvoje vzdělávání, nicméně ještě stále dost obecný pro návrh odpovídajícího hodnotícího nástroje (Standard vzdělávání ve čtyřletém gymnáziu, 1999). Tento dokument bude nahrazen Rámcovým vzdělávacím programem pro gymnaziální vzdělávání (RVP GV).

O další konkretizaci cílů i způsobu hodnocení u maturitní zkoušky se snaží katalogy požadavků, které začaly vznikat současně s úvahami o nové formě maturitní zkoušky. Jejich úkolem pro jednotlivé maturitní předměty bylo ulehčit učitelům i žákům přípravu k maturitní zkoušce. Za tímto projektem je také možno vidět snahu o stanovení určitého jednotného měřítka pro posuzování výstupních znalostí všech absolventů středních škol různého zaměření. Kromě pouhého vyjmenování požadovaných tematických okruhů se katalogy od počáteční verze z roku 2000 snažily formulovat požadované cílové kompetence (znalosti a dovednosti) a popsat formát zkoušky. Jasně a správně formulování dokumentů tohoto typu je velmi cenné a nepochybně ulehčuje práci jak žákům, tak učitelům.

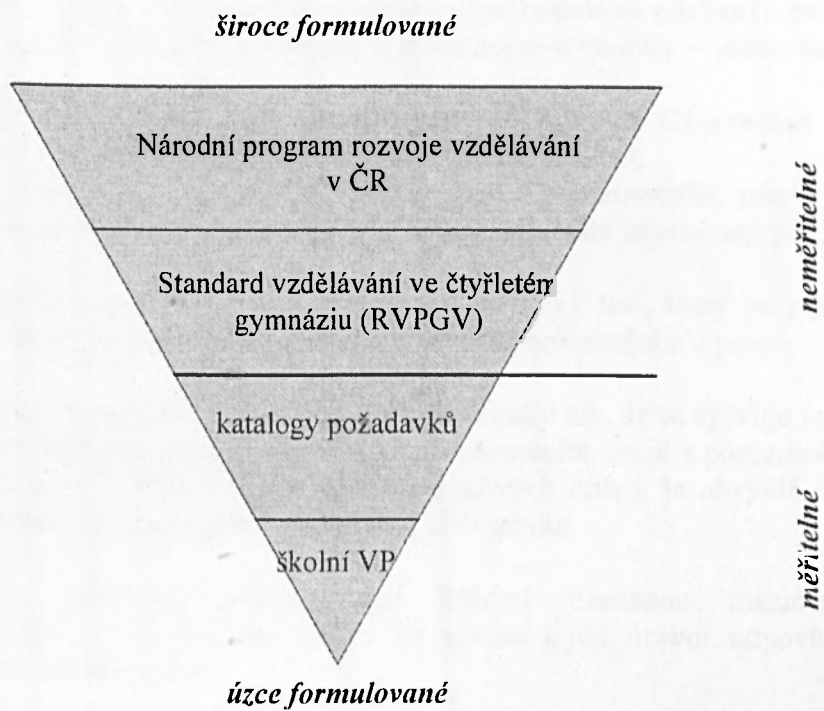
Katalogy požadavků prezentují výukové cíle jednotlivých maturitních předmětů. Informace ke katalogům požadavků, jakož i samotné katalogy požadavků pro jednotlivé maturitní předměty jsou pro zájemce dostupné na internetu. „Požadavky uvedené v katalozích jsou konkretizací vzdělávacích cílů všeobecně vzdělávacích předmětů na středních školách..... Očekávané znalosti a dovednosti jsou v katalozích konkretizovány promítnutím do tematických okruhů příslušného předmětu. Tím vzniknou vlastní maturitní požadavky (požadavky ke zkoušce zadávané MŠMT v rámci maturity). Jako synonymum je pro maturitní požadavky uveden termín specifické cíle. Maturitní požadavky jsou základní částí katalogů, představují konkrétní nároky na znalosti a dovednosti maturanta v dané oblasti učiva, které budou ověřovány příslušnou zkouškou“ (<http://www.cerimat.cz/katalogy/>). Druhá část této práce se zabývá právě hodnocením prvního Katalogu požadavků z anglického jazyka, návrhem jeho alternativní verze a také návrhem testu.

Na základě rámcových vzdělávacích programů si jednotlivé školy budou vypracovávat konkrétní školní vzdělávací programy (VP), které budou současně respektovat požadavky rámcových programů a podmínky a záměry jednotlivých škol.

Pokusme se nyní o zařazení vzdělávacích dokumentů v ČR do Muellerova grafického znázornění vzdělávacích dokumentů popisujících cíle vzdělávání (obr.1). Dokumenty se vztahují konkrétně ke vzdělávání na gymnáziích.



Obr. 8. Grafické znázornění vzdělávacích dokumentů popisujících cíle vzdělávání (Mueller, 2006, internet). Přeloženo.



Obr. 9. Zařazení vzdělávacích dokumentů v ČR do Muellerova grafického znázornění

#### 4. Testování anglického jazyka v současnosti

Jak již bylo zmíněno v subkapitole 3.2, každý testovací systém má odlišné cíle a je obvykle určen pro rozdílnou populaci, a proto není možné mluvit o univerzálním jazykovém testu, který by testoval stejné znalosti a dovednosti a používal by stejné testovací metody. Výběr odpovídajícího obsahu, formy a metod testu tedy závisí na účelu, pro který byl test navržen.

Nejčastěji jsou standardizované jazykové testy zaměřeny na zjišťování jazykové úrovně ve čtyřech základních řečových dovednostech mluvení, psaní, čtení s porozuměním a poslechu s porozuměním; pouze některé z nich zařazují také subtest zjišťující úroveň užití jazyka (*language in use* - znalost gramatických a syntaktických pravidel a slovní zásoby a schopnost jejich aplikace v určitém kontextu).

*KET (Key English Test)* je mezi cambridgeskými zkouškami angličtiny určen pro nejméně pokročilé kandidáty – podle Společného evropského referenčního rámce (dále SERR) odpovídá úrovni A2. Skládá se ze tří částí, které testují 1. čtení s porozuměním a psaní, 2. poslech s porozuměním a 3. mluvení.

*PET (Preliminary English Test)* odpovídá úrovni B1 podle SERR a testuje stejně jako *KET* čtyři základní řečové dovednosti ve třech částech: 1. čtení s porozuměním a psaní, 2. poslech s porozuměním a 3. mluvení.

*FCE (First Certificate in English)* testuje čtení s porozuměním, psaní, užití jazyka, poslech s porozuměním a mluvení a odpovídá středně pokročilé úrovni B2 podle SERR.

*IELTS (International English Language Testing System)* testuje kandidáty v uvedených dovednostech, přičemž každý z kandidátů prochází stejným testem na mluvení a poslech s porozuměním; testování čtení s porozuměním a psaní má dva moduly – jeden obecný a jeden akademický.

*CAE (Certificate in Advanced English)* odpovídá podle SERR úrovni C1 a testuje čtení s porozuměním, psaní, užití jazyka, poslech s porozuměním a mluvení.

*CPE (Certificate of Proficiency in English)* testuje čtení s porozuměním, psaní, užití jazyka, poslech s porozuměním a mluvení a podle SERR odpovídá úrovni nejvyšší, a to C2.

*TOEFL (Test of English as a Foreign Language)* je jazykový test, který se používá v USA. Testuje poslech s porozuměním, gramatiku, čtení s porozuměním a psaní.

Z uvedeného je zřejmé, že jazyková úroveň se testuje nejčastěji tak, že se zjišťuje úroveň jednotlivých řečových dovedností, tedy poslechu s porozuměním, čtení s porozuměním, mluvení a psaní. Pouze u některých testů vyšších jazykových úrovní je obvyklé, že se kromě jednotlivých řečových dovedností testuje také užití jazyka.

Následující tabulka přehledně charakterizuje každou uvedenou mezinárodní standardizovanou zkoušku z anglického jazyka ve vztahu k její úrovni odpovídající SERR a k testovaným dovednostem.

Tab. 2. Mezinárodní standardizované zkoušky z anglického jazyka, jejich úroveň podle SERR a testované dovednosti

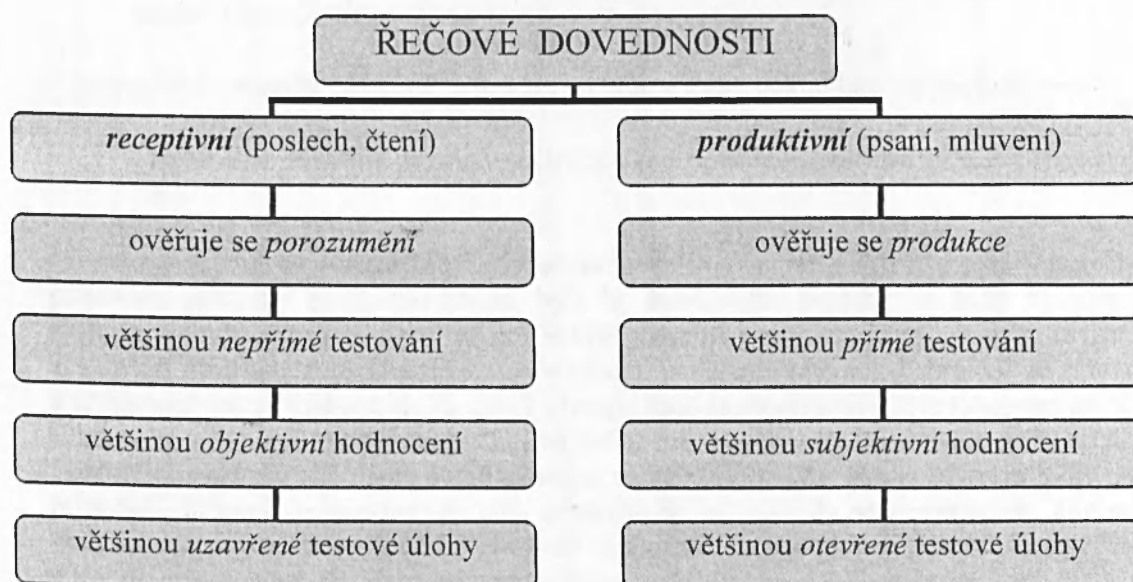
Název zkoušky	Úroveň podle SERR	Testované dovednosti
KET	A2	čtení, psaní, poslech, mluvení
PET	B1	čtení, psaní, poslech, mluvení
FCE	B2	čtení, psaní, užití jazyka, poslech, mluvení
CAE	C1	čtení, psaní, užití jazyka, poslech, mluvení
CPE	C2	čtení, psaní, užití jazyka, poslech, mluvení
IELTS	A2 – C2	poslech, čtení, psaní, mluvení
TOEFL	A2 – C2	poslech, čtení, psaní, mluvení

Pokud budeme následovat tento trend nastolený v uvedených mezinárodně uznávaných zkouškách, bylo by u středoškolské zkoušky mírně pokročilé úrovně vhodné testovat úroveň čtyř základních řečových dovedností. U středoškolské zkoušky středně pokročilé úrovně může být vhodné kromě toho také ověřování úrovně užití jazyka.

## 5. Testování základních řečových dovedností

Tato kapitola se zabývá definováním čtyř základních řečových dovedností, definováním testů těchto dovedností, obecnou charakteristikou receptivních i produktivních řečových dovedností, specifickými rysy jednotlivých řečových dovedností a dále pak nejznámějšími testovacími technikami, které se v současné době používají.

Následující obrázek znázorňuje rozdělení základních řečových dovedností a obecné charakteristiky jejich testování.



Obr. 10. Základní řečové dovednosti a charakteristiky jejich testování

### 5.1 Testování receptivních dovedností

Receptivní řečové dovednosti mají mnoho společných rysů, proto byly zařazeny do společné subkapitoly. Budou uvedeny jejich definice, charakteristika testování i nejčastěji používané testovací techniky.

#### 5.1.1 Definice receptivních dovedností

Definice dovednosti poslechu s porozuměním (dále poslech)

Pro úvahy o testování poslechu je nutné definovat, co vlastně dovednost poslechu v cizím jazyce je. *Dictionary of Language Teaching and Applied Linguistics* uvádí, že poslech je „proces porozumění řeči v druhém nebo cizím jazyce... Studium poslechu s porozuměním se při učení druhému jazyku soustřeďuje na roli jednotlivých lingvistických jednotek (např. foném, slov, gramatických struktur) jakož i na roli očekávání posluchače, roli situace a kontextu, roli obecných znalostí posluchače a roli



tématu“ (Richards; Platt; Platt 1999, s. 216). Výzkumníci se zatím neshodli na žádné jednotné definici. Jedním z uznávaných odborníků pro testování poslechu je Gary Buck, kterému byla udělena cena TOEFL již za jeho doktorskou práci v této oblasti. Od té doby, od roku 1993, publikoval mnoho dalších odborných článků a knih. Jeho definice „nastaveného/neměnného (*default*) poslechového konstruktů“ je následující:

„Je to schopnost:

- automaticky a v reálném čase zpracovat rozsáhlé vzorky realisticky mluvené řeči;
- porozumět lingvistické informaci, která je jednoznačně zahrnuta v textu a
- vyvodit jakékoliv závěry, které jsou jednoznačně implikovány obsahem pasáže textu“ (Buck, 2001, s. 114).

K jakým úvahám může uvedená definice vést tvůrce testu, pokud tuto definici přijmou?

- „Zpracovat rozsáhlé vzorky realisticky mluvené řeči, automaticky a v reálném čase.“

Je vhodné zvážit, co pokládáme za „realisticky mluvenou řeč“. Pokud by se v testech používaly nahrávky z reálného života, byly by často velmi nejasné, se šumy a jinými hluky z okolního prostředí a namluveny pravděpodobně v rychlém tempu, což by značně zvyšovalo obtížnost poslechu. Proto se v testech používají většinou nahrávky ze studia a přítomnost šumů a jiných hluků, jakož i tempo řeči, se modifikují podle toho, pro jakou jazykovou úroveň je test připravován. Pro začátečníky nebo mírně pokročilé bude určité vhodnější nahrávka téměř bez šumů a tempo řeči zřetelně nižší, než je běžné u rodilých mluvčích. Obecně se doporučuje, aby poslechový test nahrálo více mluvčích, aby se použilo více typů testových úloh i různé tematické zaměření jednotlivých testových položek, což umožní, aby vybrané vzorky řeči pokrývaly co nejširší spektrum.

Dalším problémem, který musí tvůrci testů v souvislosti s touto částí definice vyřešit je, jestli budou jednotlivé nahrávky puštěny pouze jednou nebo opakovaně – podle Bucka se má jednat o zpracování vzorku řeči „v reálném čase“. Obě možnosti (puštění nahrávky jednou nebo dvakrát) mají své obhájece; nahrávka puštěná jednou se přibližuje situaci reálného života, kdy není možné požádat o zopakování (vysílání rádia/ televize, hlášení na letišti/nádraží; přednáška). Opakovaně puštěná nahrávka zase simuluje situaci reálného života, kdy je obvykle možné v rozhovoru požádat partnera o zopakování.

- „Porozumět lingvistické informaci, která je jednoznačně zahrnuta v textu.“

Pokud má test měřit tento typ porozumění slyšenému, tvorba testových položek je nejméně problematická, protože odpovědi jsou zřejmé a jednoznačné a je jednoduché hodnotit je objektivně. Většinou jde o testové úlohy zaměřené na konkrétní, faktické otázky, týkající se údajů, které v nahrávce byly explicitně zmíněny. Zjišťování porozumění konkrétním informacím se nejčastěji ověřuje v testech, které jsou určeny kandidátům s nižší jazykovou úrovní.

- „Vyvodit jakékoliv závěry, které jsou jasně/jednoznačně implikovány obsahem pasáže textu.“

Testovat tuto schopnost kandidátů je mnohem komplikovanější než dvě předešlé. Komplikace způsobuje skutečnost, že se netestuje porozumění konkrétním a jednoznačným informacím. Naopak, kandidátovi nestačí pouhé porozumění faktům, ale musí slyšené informace dále zpracovávat; musí vyvodit nepřímo řečené ale implikované závěry ze slyšeného textu. Testové úlohy ověřující tuto schopnost je mnohem složitější a náročnější vytvořit a mnohem obtížnější je také na ně odpovědět. Testová úloha zjišťující schopnost vyvodit závěr, interpretovat určitou promluvu, tedy pochopit význam „mezi řádky“ se konstruuje mnohem složitěji než testová úloha ověřující porozumění faktické informaci. Přestože je testová položka komplikovanější, správná odpověď na ni musí být jednoznačná.

#### Definice dovednosti čtení s porozuměním (dále čtení)

Při úvahách o čtení je důležité uvědomit si rozdíl mezi čtením jako procesem a čtením jako produktem. Zkoumání čtení jako procesu je velmi náročné a komplikované. Alderson upozorňuje na to, že tento proces „se bude lišit u různých čtenářů, u různých textů, v různých časech a pro různé účely“ (Alderson, 2000, s. 3). Zkoumání procesu čtení je určitě zajímavé pro odborníky zabývající se jak vyučováním, tak i zkoušením cizích jazyků, avšak tento proces je vnitřní a probíhá u každého jedince jinak. Při tomto procesu dochází k interakci mezi textem a čtenářem; čtenář dekoduje text a odhaluje význam od nejnižších rovin grafických znaků přes slova, věty, odstavce a celé texty. Dalším možným přístupem ke čtení je jeho zkoumání jako produktu. Čtenáři pravděpodobně mohou používat nejrůznější procesy při čtení textu, avšak produkt – výsledek těchto procesů – by u různých čtenářů měl být podobný (Alderson, 2000, s. 4).

Mezi badateli zatím neexistuje shoda v definování čtení. Termíny dovednost, schopnost nebo strategie čtení se často používají ve stejném významu. *Dictionary of Language Teaching and Applied Linguistics* uvádí definici čtení jako „vnímání psaného textu za účelem porozumění jejímu obsahu“ (Richards; Platt; Platt 1999, s. 306).

Na čtení existují také různé pohledy z hlediska jeho dělitelnosti na dílčí dovednosti. Zatímco Lunzer zastává názor, že čtení je dále nedělitelnou globální dovedností, a to z toho důvodu, že existence dalších dílčích dovedností nebyla dokázána (Lunzer et al. v Alderson, 2000, s. 11), jiní odborníci se domnívají, že proces čtení zahrnuje používání různých dílčích dovedností, které se vzájemně překrývají.

Někteří z nich šli ještě dál a pokusili se jednotlivé dílčí dovednosti taxonomicky popsat. Tyto taxonomie působí logicky a jsou velmi lákavé. Například Davies uvádí osm různých dílčích dovedností čtení, Munby devatenáct a Grabe šest. Alderson však varuje, že je potřeba uvědomit si i jejich nedostatky. Uvádí tato čtyři hlavní úskalí taxonomií dovednosti čtení:

1. Taxonomie jsou založeny na teoretických úvahách spíše než na empirických pozorováních.
2. Obvykle jsou nesprávně definovány, anebo nejsou definovány vůbec.

3. Odborníci se těžko shodnou, jaké dílčí dovednosti ta která testová položka zjišťuje.
4. Analýza výkonu u testu neprokazuje existenci oddělených dílčích dovedností.  
(Alderson, 2000, s. 11).

### 5.1.2 Definice testů receptivních dovedností

#### Definice testu dovednosti poslechu

*Dictionary of Language Testing* neobsahuje definici dovednosti poslechu, pouze definici testu poslechu. Podle této publikace se jedná o „test schopnosti porozumět mluvené řeči. Poslech je proces, který nemůžeme pozorovat přímo a konstrukt poslechu s porozuměním je proto těžko postižitelný. Kromě toho kontextuální proměnné (např. počet a charakteristika mluvčích, tempo a srozumitelnost řeči, účel, téma a složitost promluvy, důvod k poslechu) ovlivňují schopnost jednotlivců pochopit mluvenou řeč“ (Davies et al., 1999, s. 110).

Z dosud uvedeného vyplývá, že jak samotná dovednost poslechu s porozuměním, tak i její měření jazykovým testem jsou velmi složité. Přesto je testování poslechu s porozuměním nedílnou součástí většiny nejznámějších standardizovaných zkoušek. Poslech patří k jedné ze základních řečových dovedností a cizojazyčnou komunikaci si bez něho nelze představit. Vandengrft uvádí, že „výzkum ukázal, že dospělí tráví 40 – 50% komunikačního času poslechem“ (Vandengrft, 2002, internet). Proto je určitě žádoucí, aby jazykové testy vyšší důležitosti obsahovaly také část zjišťující úroveň dovednosti poslechu.

#### Definice testu dovednosti čtení

*Dictionary of Language Testing* uvádí, že test dovednosti čtení je „měřítkem porozumění textu.“ Dále pak uvádí, že „testy čtení s porozuměním se pokládají za spolehlivější (protože jsou objektivní) než testy mluvení. Tradiční test čtení představuje text jako stimul a sérii testových položek s několikanásobnou volbou k měření porozumění“ (Davies et al., 1999, s. 163).

Hughes naopak nesouhlasí s názorem, že je těžší zkonstruovat spolehlivý test pro měření dovednosti mluvení nebo psaní než pro dovednosti receptivní, tedy dovednosti čtení nebo poslechu. Argumentuje tím, že zatímco když lidé mluví nebo píšou, je tento proces navenek pozorovatelný, při čtení a poslechu tento proces pozorovatelný přímo není. Proto je úkolem autorů testů „sestrojit takové testové úlohy pro čtení a poslech s porozuměním, jejichž výsledkem bude chování prokazující jejich úspěšné splnění“ (Hughes, 1992, s. 116).

### 5.1.3 Charakteristika testování receptivních dovedností

Tento oddíl se zabývá společnými i specifickými rysy testování receptivních dovedností. Uvádí také některé praktické otázky spojené s testováním čtení a poslechu.

#### Společné rysy testování receptivních dovedností

Testování dovednosti poslechu a čtení má určité společné rysy. Obě dovednosti jsou dovednostmi receptivními, avšak každá z nich je současně jedinečná a má i svá individuální specifika. Toto tvrzení podporuje také výzkum, který provedl Buck, Bae a Bachman s použitím statistických metod. Výsledky tohoto výzkumu vedly ke zjištění, že „schopnost číst a schopnost poslouchat mají hodně společného, i když každá z nich má také své individuální vlastnosti“ (Buck, 2001, s. 31).

Receptivní dovednosti lze měřit s použitím testů nepřímých. Dovednost čtení a poslechu není totiž možné přímo pozorovat a hodnotit; čtení a poslech jsou procesy, které probíhají vnitřně a nejsou navenek pozorovatelné.

Podobnost existuje i při zpracovávání slyšeného nebo čteného textu. U obou řečových dovedností se hovoří o dvou procesech, a to o procesu „zdola nahoru“ (*bottom-up*) a o procesu „shora dolů“ (*top-down*).

Proces „zdola nahoru“ při zpracovávání *poslechového materiálu* předpokládá, že posluchač nejdříve identifikuje fonémy, slova, a postupně věty a vyšší celky promluvy, zpracovává sémantický obsah, dostává se k doslovnému významu a poté, na základě konkrétní komunikační situace, ke skutečnému významu promluvy mluvčího. Tento proces vypadá logicky, avšak pravděpodobně neprobíhá v těchto za sebou následujících stádiích.

Naopak, porozumění slyšenému pravděpodobně probíhá v procesu „shora dolů“, což znamená, že zpracovávání slyšeného textu probíhá na různých rovinách simultánně nebo v jakémkoliv vhodném pořadí. Při porozumění slyšenému nám pak nepomáhají pouze znalosti lingvistické, ale jakékoliv jiné znalosti o životě, světě kolem nás, o tématu nebo kontextu mluveného projevu (Buck, 2001, s. 2-3).

Model „zdola nahoru“ při procesu *porozumění čtenému* předpokládá, že čtenář musí při procesu čtení začínat tištěným slovem, přičemž rozpoznává grafické stimuly, proměňuje je na slova a dekóduje jejich význam. Tento model je spojován s teorií behaviorismu (Alderson, 2000, s. 16-17).

V protikladu k tomuto přístupu stojí model „shora dolů“ vycházející z teorie schémat, která předpokládá, že text je zpracováván pomocí aktivace schémat. „Čtenáři aktivují ta existující schémata, která pokládají za relevantní, a přicházející informaci do nich založí. Čtení je pak úspěšné do té míry, jak jsou tato schémata relevantní“ (Alderson, 2000, s.17).

Společné rysy dovedností poslechu a čtení se týkají i praktických otázek spojených s jejich testováním. Například při hodnocení testů poslechu i čtení se doporučuje, aby v případě produktivních odpovědí nebyli kandidáti penalizováni za možné gramatické nebo pravopisní chyby, pokud je zjevné, že mluvený nebo psaný test pochopili. Pokud bychom při hodnocení poslechu nebo čtení chtěli současně hodnotit i úroveň produktivní dovednosti psaní, bylo by měření poslechu nebo čtení méně přesné (Hughes, 1992, s. 131). Další možností, jak se vyhnout kontaminaci hodnocení dovednosti poslechu nebo čtení hodnocením dovednosti psaní je umožnit kandidátům odpovědět na otázky v mateřském jazyce. Tento postup je samozřejmě možný pouze u skupin kandidátů se stejným mateřským jazykem, který je také mateřským jazykem examinatorů.

Při testování obou receptivních dovedností se také doporučuje, aby jak nahrávky, tak texty byly vždy zasazeny do určitého kontextu; v praxi to znamená, že je vhodné, aby jednotlivé testové úlohy obsahovaly krátký nadpis, uvádějící kandidáty do situace, ve které se slyšený nebo psaný text vyskytuje.

Při tvorbě položek testu zjišťujícího úroveň poslechu a čtení se doporučuje, aby byly položky prezentovány v takovém pořadí, jak se objevují v nahrávce/textu. Dále je vhodné, aby odpovědi na jednotlivé položky nebylo možné správně vypracovat i bez porozumění slyšenému/čtenému, například na základě faktických znalostí. Použité texty by měly být pro cílovou populaci zajímavé, neměly by na cílovou populaci působit urážlivě a neměly by být vybírány ze žádných výukových materiálů.

#### Specifické rysy testování dovednosti poslechu

Při úvahách o testování poslechu nesmíme zapomínat, že existuje velký rozdíl mezi jazykem psaným a mluveným. Tento fakt je klíčový při výběru vhodných textů a v praxi znamená, že pro testování poslechu není vhodné používat texty určené pro čtení. Mluvený text je obvykle kratší a jednodušší než text psaný. V mluvené řeči se často používá jednodušších syntaktických struktur, kratších vět a hustota informací je obvykle nižší než u psaného textu. Mluvíci při ústní produkci používají gesta, zkracují věty, řeč je často rozdělována pauzami, některé výrazy jsou mluvcím opakovány (Buck, 2001, s. 10-11). Poslechový test je zadáván takovým způsobem, že kandidát musí na slyšený text reagovat současně s poslechem nahrávky, resp. ihned po jejím vyslechnutí. Ve srovnání s vypracováváním odpovědí u testu čtení je vypracování odpovědí u testu poslechu procesem daleko náročnějším. Zatímco u testu čtení je možné text přečíst několikrát anebo se k určitým částem opakovaně vrátit, v testu poslechu to možné není. Poslech nahrávky probíhá v reálném čase a stejným tempem pro všechny kandidáty, kdežto pro test čtení je sice většinou stanoven stejný časový limit, ale tempo a pořadí vypracovávání odpovědí si určuje každý kandidát sám.

Buck rozlišuje dva základní druhy znalostí, které jsou potřebné pro porozumění slyšenému – lingvistické a nelingvistické, přičemž lingvistické znalosti zahrnují znalosti fonologie, slovní zásoby, sémantiky a syntaxe a nelingvistické znalosti představují všeobecné znalosti o životě a světě, znalosti kontextu, ve kterém se poslechová situace

odehrává, a samozřejmě také znalosti konkrétní problematiky, o které poslechový materiál pojednává (Buck, 2001, s. 247).

#### Praktické otázky spojené s testováním dovednosti poslechu

Z odborné literatury lze pro testování poslechu vyabstrahovat určitá doporučení. Nevycházejí vždy pouze z výsledků výzkumu; častěji znamenají spíše současný trend v jazykovém testování. Uvedený seznam doporučení rozhodně nezahrnuje všechna doporučení, ale pouze ta nefrekventovanější:

- Poslech se nejčastěji testuje v jazykových laboratořích, i když v některých případech se hodnotí také nebo pouze v průběhu ústní zkoušky (tyto případy jsou však spíše ojedinělé).
- Kandidáti mají nejdříve možnost alespoň rychlého přečtení testových úloh; tento postup doporučuje většina odborníků, přičemž jedním z důvodů je skutečnost, že v reálném životě je dovednost poslechu potřebná vždy v konkrétní situaci a v konkrétním kontextu, který je mluvčím známý. Pokud by při testování poslechu byly testové úlohy předloženy kandidátům až po ukončení nahrávky, jednalo by se o situaci poněkud nepřírozenou. U většiny známých standardizovaných testů se postupuje tak, že kandidáti dostanou testové úlohy před samotným poslechem, i když výzkumy jednoznačně nepotvrdily, že možnost seznámit se s otázkami předem přináší lepší výsledky kandidátů (Buck, 2001, s. 137).
- Testové položky by měly být formulovány stručně a jasně, aby nevyžadovaly mnoho čtení.
- Každá testová položka by se měla řešit nezávisle; tedy správné vypracování jedné položky by nemělo být závislé na správném vypracování jiné položky.

Při tvorbě testu a testování poslechu je potřeba brát v úvahu mnohé skutečnosti. Tvůrci testů musí hledat odpovědi na různé konkrétní otázky, týkající se používání cizího jazyka kandidáty, např.:

- Budou kandidáti v reálném životě komunikovat s rodilými mluvčími hovořícími britskou, americkou, australskou angličtinou a/nebo s nerodilými mluvčími angličtiny?
- Budou poslouchat většinou monology nebo dialogy?
- Budou vystaveni cizímu jazyku v přímém rozhovoru (tváří v tvář), telefonicky, jazyku z amplionů na letišti, na nádraží, nebo jazyku médií (televize, rádio, film, video)?
- Jaké bude prostředí, ve kterém budou komunikovat (tiché/hlučné, např. letiště, hotel, letadlo, vojenský štáb, letištní věž, nemocnice)?
- S jakou obtížností poslechových materiálů se budou nejčastěji setkávat v reálném životě?
- Pokud se budou pohybovat v pracovním prostředí, jaké dovednosti jejich profese vyžaduje z hlediska jazyka obecně a poslechu konkrétně?

Odpovědi na uvedené otázky by měli tvůrci testů zjistit předtím, než začnou o testu jako takovém přemýšlet. Někdy je možné potřebné informace získat relativně jednoduchým způsobem, a to v případě, že test bude určen jasně charakterizované, nejčastěji profesně jednotné skupině kandidátů a používání cizího jazyka bude poměrně detailně vymezeno, např. test pro stíhací piloty, lékaře, hotelové zaměstnance, řidičí letového provozu apod. Častěji se však jedná o testy, kterými budou testovány daleko různorodější skupiny kandidátů. V takových případech je potřeba identifikovat reprezentativní vzorek co možná nejširšího spektra situací, v nichž se bude kandidát nacházet, a test navrhnout tomu odpovídajícím způsobem.

Pokud autoři testu najdou spolehlivé odpovědi na uvedené otázky, budou se dále rozhodovat jako jazykoví odborníci o záležitostech, jako například:

- Bude se hodnotit porozumění detailu, jednomu slovu, celkovému významu, hlavní myšlence?
- Bude se testovat pouze poslech s porozuměním, nebo i faktické znalosti, případně paměť?
- Mají být poslechové testy autentické nebo upravené?
- Mají být poslechové testy předneseny pouze rodilými nebo i nerodilými mluvčími?
- Mají být poslechové testy nahrány a reprodukovány do sluchátek nebo čteny ve třídě?
- Je vhodné, aby byl poslechový test nahrán ve studiu a zcela bez šumů nebo v reálném prostředí a se šumy?
- Budou kandidáti slyšet nahrávku jednou nebo opakovaně?
- Jaké má být tempo řeči v nahrávce?

Když tvůrci testů najdou odpovědi i na tyto otázky, budou se rozhodovat, jakou testovací techniku, resp. jaké testovací techniky zvolí.

#### Specifické rysy testování dovednosti čtení

Heaton se domnívá, že při testování čtení je důležité, aby ověřované dovednosti v druhém jazyce měli kandidáti již zvládnuté v prvním (mateřském) jazyce (Heaton, 1991, s. 105). Další zajímavou myšlenkou Heatonova je jeho doporučení používat v testech čtení různorodé autentické texty z reálného života. Domnívá se, že to nejenom poskytne realističtější prostředky hodnocení, ale že to také bude kandidáty motivovat (Heaton, 1991, s. 107).

#### Praktické otázky spojené s testováním dovednosti čtení

Při testování čtení je důležité zvážit různé skutečnosti. Budou se samozřejmě lišit od těch, které tvůrci testu zvažují u testování jiných řečových dovedností. V případě testování čtení s porozuměním budou tvůrci testů zajímat otázky, jako například:

- Jaké typy textů budou kandidáti v reálném životě číst (jednoduché vzkazy, formální nebo neformální dopisy, informační a turistické brožury, inzeráty, jízdní řády, formuláře, články v novinách a časopisech)?
- V jakých situacích a s jakým záměrem budou číst?
- Bude se jednat o odborné texty z určité oblasti?
- Jak rozsáhlé budou tyto texty?
- Jak obtížné budou tyto texty?
- Bude úkolem kandidátů přečtené texty nějakým způsobem dále zpracovávat (shrnutí textu, výběr hlavních myšlenek apod.)?

V závislosti na účelu, pro který má být test vytvořen, je někdy možné od očekávaných kandidátů nebo jejich nadřízených tyto informace získat neformálním způsobem, někdy je zase vhodné požádat o informace písemně formou dotazníku. V jiných případech naopak může být skupina kandidátů různorodého zaměření a informace od nich získané se mohou velmi lišit. V takovém případě je na tvůrcích testu, aby odpovědně zvážili výběr textů pro test.

Poté co se identifikují typy textů, jejich obtížnost a případně i tematické zaměření, je nutné rozhodnout o tom, jaké dílčí dovednosti čtení by test měl měřit. Autoři testů budou tedy řešit otázky, jako například:

- Bude se hodnotit porozumění detailu, jednomu slovu, celkovému významu, hlavní myšlence?
- Bude se testovat pouze čtení s porozuměním, nebo i faktické znalosti, případně paměť?

V neposlední řadě stojí před tvůrci testu rozhodnutí, do jaké míry mají být texty pro čtení s porozuměním autentické nebo upravené.

#### 5.1.4 Testovací techniky receptivních dovedností

Obecně se doporučuje používat spíše více testovacích metod než pouze jednu: „nejlepší radou, kterou je možné doporučit pisatelům testových položek je, aby používali více než jednu testovací metodu pro testování jakékoliv z dovedností“ (Alderson; Clapham; Wall 1995, s. 45). Výzkum dále také ukazuje, že testovací techniky ovlivňují výsledek testu. Ve skutečnosti pravděpodobně neexistuje testovací technika, která by neměla žádný vliv na výsledek testu. To znamená, že vliv testovacích metod nemůžeme vyloučit, a proto je potřebné se s charakteristikami testovacích metod seznámit a snažit se jim porozumět (Bachman; Palmer 1996, s. 46).

Testovací techniky společné pro receptivní řečové dovednosti

Výběr odpovědi z vícenásobné nabídky (*multiple choice*)



Tato metoda je používána často a nabízí kandidátovi odpovědi většinou formou možností a, b, c nebo d. Může způsobit problémy kandidátům, kteří na tento způsob testování nejsou zvyklí. Na druhou stranu má také své výhody, a to zejména úplně objektivní, spolehlivé a rychlé hodnocení odpovědi. Jednotlivé položky je snadné statisticky vyhodnocovat. Tento způsob testování je u celé řady odborníků velmi oblíbený, a to zejména v oblasti velkoplošného testování (*large-scale testing*) zjevně z praktických důvodů. Avšak laici i učitelé jazyků jsou často vůči této metodě poměrně skeptičtí.

Psaní položek je, přestože to tak na první pohled nevypadá, poměrně obtížné, protože není jednoduché vytvořit správně otázku, jednu správnou odpověď a také tři distraktory, které by měly být přibližně stejně dlouhé a žádný z nich by neměl být lehce vylučitelný. Správná odpověď by se neměla odlišovat od tří distraktorů rozdílnou délkou nebo strukturou.

Hughes uvádí několik hlavních nevýhod tohoto způsobu testování:

- „Tato technika testuje pouze rozpoznávací znalost.
- Hádání může mít značný, ale nerozpoznatelný vliv na výsledky testu.
- Technika silně omezuje rozsah toho, co se může s jejím použitím testovat.
- Je velmi těžké napsat vhodné/dobré položky.
- Zpětný efekt na vyučování může být škodlivý.
- Tato technika může ulehčit podvádění.“ (Hughes, 1992, s. 60-62)

Je možné souhlasit s jeho kritikou této testovací techniky?

- „Tato technika testuje pouze rozpoznávací znalost.“

S tímto tvrzením se do určité míry dá souhlasit, protože tato technika opravdu nabízí kandidátům „hotové“ odpovědi a tudíž pro úspěšné vyřešení testové úlohy stačí, aby kandidát správnou odpověď rozpoznal od nesprávných (distraktorů). Tato technika tedy nevyžaduje od kandidáta žádnou produkci, vytvoření psaného nebo ústního projevu, což je na druhé straně výhodou, protože vylučuje kontaminaci hodnocení receptivních dovedností hodnocením dovedností produktivních.

V případě dobře zkonstruovaných testových položek testujících zejména vyšší úroveň řečových dovedností však s Hughesem nelze zcela souhlasit, protože se pak nejedná pouze o rozpoznání správné odpovědi. Aby kandidát rozpoznal správnou odpověď, musí jazykový materiál nejdříve zpracovat, interpretovat a v některých případech i takzvané „číst mezi řádky“. Buck se domnívá, že tento typ testových úloh „může být použit k testování různých dílčích poslechových dovedností: od porozumění na nejexplicitnější doslovné úrovni, přes kombinování informací z různých částí textu, ...porozumění implicitních významů až k sumarizaci a syntéze rozsáhlejších pasáží“ (Buck, 2001, s. 146).

- „Hádání může mít značný, ale nerozpoznatelný vliv na výsledky testu.“

Podle toho, kolik možností v testu se nabízí, je pravděpodobnost uhodnutí správné odpovědi přibližně 33% (s nabízenými možnostmi a, b, c) a 25% (s nabízenými možnostmi a, b, c, d). Kandidáti, kteří se snaží uhodnout správnou odpověď, mohou mít počet bodů vyšší nebo nižší, než odpovídá jejich znalostem. Wu ve svém výzkumu došel k závěru, že tato technika zvýhodňuje kandidáty na pokročilé úrovni a že „nesprávné interpretace voleb vedly některé kandidáty k výběru nesprávné možnosti a naopak, někteří kandidáti vybrali správnou odpověď, ale z nesprávných důvodů“ (Buck, 2001, s. 143).

- „Technika silně omezuje rozsah toho, co se může s jejím použitím testovat.“

Tato technika zjevně nemůže být použita k testování psaní nebo mluvení, tedy produktivních řečových dovedností. Kromě toho Hughes argumentuje skutečností, že pro vytvoření testových položek s výběrem odpovědi je často problematické najít vhodné distraktory. Jako příklad uvádí některé gramatické jevy, u kterých příprava distraktorů může být problematická. (Hughes, 1992, s. 61). V tomto bodu lze s Hughesem souhlasit.

- „Je velmi těžké napsat vhodné/dobré položky.“

Přestože testové položky s výběrem odpovědi vypadají na první pohled velmi jednoduše, je opravdu složité vytvořit dobrou, správně fungující testovou položku. „Konstruování těchto položek je vysoce profesionální dovednost, jejíž osvojení si vyžaduje hodně času a praxe“ (Buck, 2001, s. 142). S tímto názorem souhlasí i Chráska a dodává, že při tvorbě distraktorů je možné postupovat tak, „že se úloha nejdříve zadá žákům jako otevřená a potom se nejčastěji vyskytujících chyb použije jako distraktorů“ (Chráska, 1999, s. 35).

Ve srovnání s kritickými připomínkami Hughese kolektiv autorů *Dictionary of Language Testing* uvádí také kladné stránky této testovací techniky. K nim podle autorů patří zejména jednoduché vyhodnocování položek, zjištění jejich obtížnosti a reliability jakož i reliability celého testu. Kromě toho je s použitím této testovací techniky možné zahrnout do jednoho testu více různých vzorků z oblasti, kterou chceme testovat (gramatické jevy, tématické oblasti apod.). Autoři této publikace však zároveň zdůrazňují potřebu pečlivého připomínkování, pretestování a následných úprav tohoto typu položek. Za silný argument, proč tuto techniku používat, možno také pokládat skutečnost, že „položky s výběrem odpovědi se používají ve velkém měřítku, zvláště v standardizovaných testech, déle než posledních 50 let“ (Davies et al., 1999, s. 125). Dalším pozitivním rysem této testovací techniky, který publikace zmiňuje, je zjištění vysokých korelací mezi výsledky testů s výběrem odpovědí a výsledky testů produktivních dovedností. Tato skutečnost poskytuje důkaz, že technika výběru odpovědi zjišťuje úroveň určitých řečových schopností, které jsou společné pro různé druhy používání jazyka (Davies et al., 1999, s. 125). Kromě výhod tohoto způsobu testování vidí autoři této publikace také jeho negativní vlastnosti. Podobně jako Buck připouštějí, že tato testovací technika dává kandidátům možnost pouze rozpoznat správnou odpověď a ne produkovat jazyk. Uvádějí názor, že „základní charakteristický rys, který je potřebný pro správnou odpověď, není znalost jazyka, ale nějaký jiný, irelevantní konstrukt, jako například schopnost vyloučit zjevně nesprávné distraktory. Toto tvrzení je podpořeno

studiemi, které ukazují, že kandidáti mohou své výsledky u testu s výběrem odpovědi značně zlepšit tím, že si je opakovaně procvičují“ (Davies et al., 1999, s. 125).

Přestože je technika výběru odpovědi praktická pro statistické vyhodnocování, nesmíme zapomínat, že jí chybí autenticita. Není možné představit si situaci reálného života, ve které by kandidáti byli postaveni před podobnou úlohu.

Další ještě nezmíněnou negativní vlastností této testovací metody může v některých případech být skutečnost, že formou distraktorů nabízí kandidátům často nesprávné odpovědi, které by je vůbec nenapadly a tudíž je mohou mást. Někteří badatelé jdou v kritice této metody tak daleko, že tvrdí, že schopnost odpovědět na tento typ testových položek je zvláštní schopnost, která je odlišná od schopnosti číst s porozuměním (Alderson, 2000, s. 211).

Pokud mají být testové položky vytvořeny touto metodou kvalitní, je při jejich tvorbě nepostradatelná pomoc kolegů, resp. rodilých mluvčích, kteří na testové položky sami odpovědi a kriticky se k nim vyjádří. Poté musí být položky pretestovány na vhodné populaci a vyhodnoceny. Nevhodné položky je pak potřeba buď z testu odstranit, nebo je přepracovat. Proto se doporučuje vytvořit vždy o něco větší počet položek, než má být použito v testu.

Heaton upozorňuje, že při tvorbě položek je důležité vyvarovat se toho, aby kandidáti mohli víceméně přiřadit správnou odpověď (jednu z nabízených voleb, nejčastěji a, b, c nebo d) k části textu. Proto doporučuje, aby položky požadovaly od kandidátů „shrnout, utřídit a interpretovat, co přečetli“ (Heaton, 1991, s. 119).

Alderson uznává, že tato metoda se při testování používá často, avšak zároveň upozorňuje, že „by bylo naivní předpokládat, že metoda je „validní“ právě proto, že se používá ve velkém rozsahu“ (Alderson, 2000, s. 204). Hughes přisuzuje této metodě „povrchní přitažlivost“ (Hughes, 1992, s. 120).

Z uvedených názorů odborníků na tuto testovací techniku je patrné, že přes některé negativní vlastnosti také disponuje mnohými klady. Pokud jsou tvůrci testů pro tvorbu tohoto typu položek vyškoleni, je vhodné tuto techniku při testování receptivních dovedností použít.

Pro tvorbu testových úloh výběru odpovědi s vícenásobnou volbou platí několik doporučení:

- Měla by být správná vždy pouze jedna volba.
- Všechny nabízené volby by měly být gramaticky správné.
- Všechny odpovědi by měly být přibližně stejné délky i obtížnosti.
- Všechny odpovědi by měly mít stejnou konstrukci (stejný slovní druh).
- V zadání by se neměl používat zápor.
- Distraktory by měly přitáhnout pozornost kandidátů, kteří si nejsou jisti správnou odpovědí.

- „Distraktory, které si vybere málo kandidátů, přispívají položce jenom málo, protože účinně snižují počet alternativ a tím zvyšují pravděpodobnost, že kandidáti dospějí ke správné odpovědi pouze hádáním“ (Davies et al., 1999, s. 49).
- Distraktory by neměly být absurdní, naopak měly by vypadat přijatelně.

#### Pravdivá a nepravdivá tvrzení (*true – false*), dichotomní položky (*dichotomous items*)

Princip této techniky je jednoduchý. Po poslechu testové nahrávky nebo po přečtení textu se má kandidát rozhodnout, jestli jsou uvedená tvrzení v souladu s tím, co slyšel nebo četl (pravdivá), nebo v protikladu s tím, co slyšel nebo četl (nepravdivá).

Za nevýhodu této testovací techniky se nejčastěji pokládá skutečnost, že kandidáti mohou správnou odpověď uhodnout, přestože obsah slyšeného či přečteného textu správně nepochopili. Pravděpodobnost správného uhodnutí je ještě vyšší než u techniky výběru odpovědi z vícenásobné nabídky – 50%. Riziko uhodnutí je možné snížit buď tím, že kromě možnosti pravdivé – nepravdivé tvrzení může být uvedena možnost „nebylo zmíněno.“ Dopad uhodnutí správné odpovědi je možné zredukovat také tím, že zvýšíme počet testových položek. Jiný názor zastávají autoři *Dictionary of Language Testing*. Pro omezení hádání správných odpovědí doporučují, aby se za nesprávné odpovědi strhávaly body (Davies et al., 1999, s. 215).

Hughes tuto metodu pokládá za variantu metody výběru odpovědi z nabídky s pouze jedním distraktorem (Hughes, 1992, s. 121). Heaton oceňuje jednoduchost a rychlost hodnocení a domnívá se, že výsledky poskytují spolehlivé ukazatele úrovně čtení s porozuměním, pokud jsou položky správně navrženy a je jich dostatečný počet (Heaton, 1991, s. 113). Kromě toho však upozorňuje, že test s položkami tohoto typu může špatně rozlišovat mezi kandidáty, a to zejména v případě, že není vytvořen dostatečný počet položek (Heaton, 1991, s. 113).

#### Otázky s krátkými odpověďmi (*short answer questions*)

Pro testování poslechu i čtení je tato technika vhodná. Vyžaduje od kandidáta, aby stručně odpověděl na danou otázku. Požadovaná odpověď má být krátká, aby bylo možné zajistit objektivní a spolehlivé hodnocení. Jediný způsob, kterým je možné tohoto docílit, je položky pretestovat a zjistit, jestli očekávaná odpověď je jedinou správnou odpovědí. Pokud pretestování odhalí další možné správné odpovědi, je potřeba buď položku přepracovat, nebo tyto další možné odpovědi zařadit do klíče. Ve srovnání s metodou výběru odpovědi z nabídky nebo s metodou pravdivých a nepravdivých tvrzení při technice krátkých odpovědí kandidát sám aktivně prezentuje své porozumění, resp. neporozumění textu, což je možné pokládat za výhodu této metody.

Vytvoření takovýchto otázek nebývá složité. Variantou této techniky může být například doplňování vět. Hughes tento typ testových úloh hodnotí kladně: „Varianta (krátkých odpovědí) s doplňováním, která vyžaduje od kandidáta minimum psaní, je obzvláště užitečná“ (Hughes, 1992, s. 137).

### Informační transfer (*information transfer*)

Tato technika v praxi znamená, že informace, které kandidát slyší v mluveném textu nebo přečte v psaném textu, má zaznačit do plánu, mapy, formuláře, tabulky apod. Jejím kladem je, že od kandidátů vyžaduje minimální produkci psaného projevu, čímž přispívá k objektivnímu hodnocení odpovědí.

O této technice lze říci, že je autentická, protože se kandidáti před podobnými úkoly mohou v reálném životě ocitnout. Na druhé straně tento typ testové úlohy, pokud použije plánek určitého místa nebo mapu, může znevýhodnit kandidáty, kteří se neumějí orientovat v mapě (Alderson; Clapham; Wall 1995, s. 52).

### Doplňování chybějících informací (*gap-filling, completion items*)

Při použití této techniky mají kandidáti k dispozici text, ve kterém chybějí některá slova nebo výrazy. Někdy mají k doplnění chybějících informací seznam slov a/nebo výrazů, které mají k doplnění použít; v tom případě je potřebné dát k dispozici více slov/výrazů než je mezer. Jindy takovýto seznam k dispozici nemají a slova a/nebo výrazy mají doplňovat pouze podle obsahu nahrávky či textu.

Chybějící slova jsou vybírána racionálně, tedy podle uvážení pisatele testu. Mezi dvěma mezerami by nemělo být ponecháno méně než pět až šest slov, aby se kandidáti mohli v textu orientovat. Tento typ testovací techniky označuje Heaton jako doplňovací položky (*completion items*), (Heaton, 1991, s. 124).

Nejčastějším problémem tohoto typu testových úloh je skutečnost, že do každé mezery by kandidáti měli napsat pouze jedno slovo. Toto slovo by mělo být stejné u všech kandidátů. V praxi je splnění tohoto požadavku často velmi problematické, a proto je v klíči k testu obvykle uvedeno více možností, které jsou přijatelné jako správné odpovědi. Z tohoto důvodu je velmi důležité položky před „ostrým“ použitím v testu pretestovat – pouze při pretestování se často ukáže, že odpověď není tak jednoznačná, jak by se na první pohled mohlo zdát, a že existuje více možných odpovědí, které se pak doplní do klíče.

### Shrnutí (*summary*)

Kandidáti si vyslechnou nahrávku nebo přečtou text a poté napíší nebo ústně zreprodukují shrnutí nahrávky či textu. V tomto případě, na rozdíl například od metody krátkých odpovědí, není možné hodnotit separátně pouze poslech nebo čtení s porozuměním. Musí se současně hodnotit také buď psaní nebo mluvení. Proto tuto metodu hodnocení není možné pokládat za objektivní. Aby však hodnocení shrnutí bylo co nejvíce objektivní, je velmi důležité navrhnout jasná a srozumitelná kritéria hodnocení. Při hodnocení totiž není možné „uvádět úroveň receptivní dovednosti odděleně od úrovně produktivní dovednosti, protože není jasné, která z těchto dovedností stojí za silnými i slabými stránkami výkonu“ (Davies et al., 1999, s. 192).

Přestože hodnocení může být složitější, je možné tuto metodu pokládat za autentickou. Pokud mají všichni kandidáti i hodnotitelé stejný mateřský jazyk, je jednou z možností, jak odstranit z bodování potřebu hodnotit také písemný projev, požadovat shrnutí textu v mateřském jazyce. Alderson doporučuje řešit problém kontaminace hodnocení čtení hodnocením psaní tím, že se kandidátům dá k dispozici několik shrnutí, z nichž mají vybrat to nejlepší (*multiple-choice summaries*), (Alderson, 2000, s. 236).

Alderson se dále domnívá, že pro vytvoření shrnutí musí kandidát pochopit hlavní myšlenky. Tato skutečnost naráží na problém, protože identifikování hlavních myšlenek textu může být velmi subjektivní. Jako řešení navrhuje, aby pisatelé testu napsali nejdříve svá shrnutí a aby jako přijatelná shrnutí byla určena ta, která obsahují hlavní myšlenky, na kterých se shodli všichni nebo alespoň většina tvůrců testu (Alderson, 2000, s. 232).

Shrnutí *cloze*, shrnutí s doplňováním mezer (*summary cloze, gapped summary*)

Jako vhodnou variantu shrnutí doporučuje Alderson shrnutí s doplňováním mezer (Alderson, 2000, s. 240). Kandidáti dostanou shrnutí přečteného textu, ve kterém chybějí některá slova. Úkolem kandidátů je doplnit jedno nebo dvě slova, a to buď samostatně, nebo je možné poskytnout jim seznam slov, ze kterého text doplňují. Stejnou techniku je možné použít pro testování poslechu.

Pokud má kandidát do jedné mezery doplnit více než jedno nebo dvě slova, bude kromě dovednosti porozumění prezentovat také dovednost psaní. „Je sporné, co se vlastně touto formou testu testuje“ (Davies et al., 1999, s. 192).

Pokud všichni kandidáti i examinátoři mluví stejným jazykem, je také možné doplňovat slova v rodném jazyce. Použitím této techniky se odstraní komplikované hodnocení shrnutí, které píší kandidáti v testovaném jazyce.

Přiřazování (*matching*)

Při této testovací metodě mají kandidáti přiřadit k sobě například dvě skupiny slov, které spolu souvisejí, definice slov ke slovům, nadpisy k odstavcům nebo textům, věty nebo fráze k obrázkům, popisy nebo vyprávění příběhů k obrázkům a podobně. Při použití této metody je důležité, aby počet položek ve skupinách, které se vzájemně přiřazují, nebyl stejný např. kandidáti budou mít čtyři různé texty, ale pět anebo více nadpisů, které k textům mají přiřadit. V opačném případě by poslední odpověď mohla být správně vypracována i bez správného porozumění.

Seřazování (*ordering, rearrangement, identifying order, jumbled text*)

Při této testovací metodě jsou kandidátům prezentovány věty, části textů nebo obrázky znázorňující obsah textu v přeházeném pořadí. Úkolem kandidátů je tyto prvky uspořádat do správného pořadí na základě porozumění přečtenému nebo slyšenému textu. Opet však stojí za připomenutí, že tuto testovací metodu lze jen stěží pokládat za autentickou. Jistý problém může nastat také při hodnocení, protože není jasné, jestli se má hodnotit

zvláště každé individuální přiřazení nebo pouze testová úloha jako celek. Alderson pokládá metodu přiřazování a seřazování za „alternativní objektivní techniku“ (Alderson, 2000, s. 215, 219).

Testovací techniky specifické pro testování dovednosti poslechu

Zapisování poznámek (*note-taking*)

Další technikou zjišťující úroveň poslechu je zapisování si poznámek podle slyšeného textu. Tento typ testových úloh je vhodné zařadit do testu jenom v takovém případě, když se předpokládá, že kandidáti budou tuto aktivitu vykonávat v reálném životě; jedná se například o kandidáty, kteří budou studovat na univerzitě nebo v kurzech v zahraničí.

Při aplikaci této techniky kandidáti poslouchají mluvený text, dělají si poznámky a až poté, na základě zaznamenaných poznámek, odpovídají na otázky (písemně nebo ústně). Pokud by se totiž hodnotily pouze poznámky, bylo by toto hodnocení velmi subjektivní. Proto stojí za uvážení, jestli se bude hodnotit pouze úroveň dovednosti poslechu, nebo také úroveň psaní či mluvení.

Testovací techniky specifické pro testování dovednosti čtení

Doplňování *cloze*

Při použití této testovací metody autor testu odstraní každé n-té slovo v textu a nahradí ho mezerami. Úkolem kandidátů je tyto mezery doplnit. Na rozdíl od doplňování chybějících informací mezery nevolí dle vlastního uvážení tvůrce testu, ale vynechává se nejčastěji každé páté, šesté, sedmé, osmé, deváté nebo desáté slovo. Běžně se slova začínají vynechávat až po první větě. Kdyby se vynechalo každé čtvrté slovo, kandidáti by mohli mít velké potíže s pochopením textu. Tato metoda se používá jednak k měření úrovně dovednosti čtení, někdy však také k měření celkové jazykové schopnosti. K bodování existují obecně dva hlavní přístupy: buď se bod uděluje pouze za slovo/výraz, které bylo v původním textu, nebo i za doplnění jakéhokoliv jiného vhodného slova/výrazu.

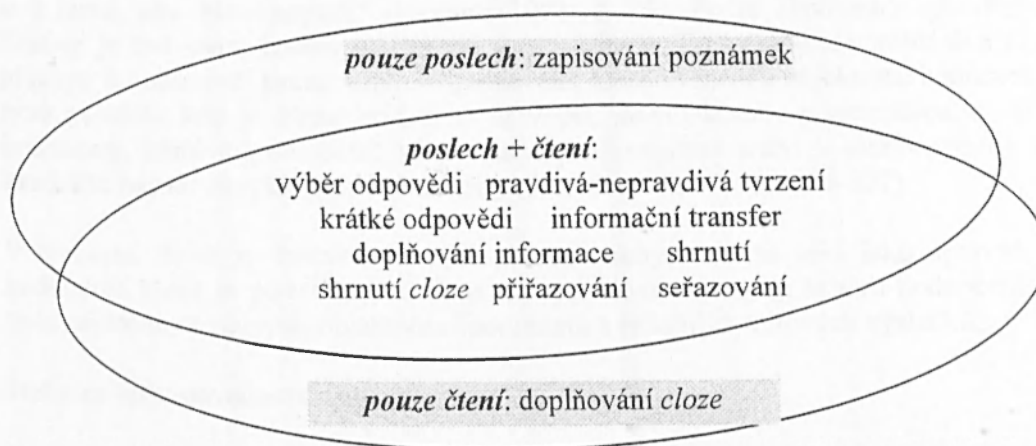
V minulosti existovalo období, kdy tato metoda byla velmi oblíbenou a předpokládalo se, že měří určitou globální jazykovou schopnost. Avšak ze zkušenosti víme, že většina jedinců nemá jednotlivé řečové dovednosti na stejné úrovni. Je možné, aby stejná osoba v cizím jazyce například dobře psala, ale přitom měla problémy s ústním vyjadřováním. Proto pravděpodobně není správné uvažovat o celkové jazykové schopnosti; většinou nás zajímá, jakou úroveň kandidáti mají v jednotlivých řečových dovednostech. Kromě toho se tato metoda také nejeví jako metoda autentická; podobné úkoly na kandidáty v reálném životě nečekají.

Další skutečností, která vzbuzuje jisté pochybnosti o vhodnosti použití této metody je, že i vzdělání rodilí mluvčí se vzájemně podstatně liší ve schopnosti předpovědět, jaké slovo bylo z textu odstraněno (Hughes, 1992, s. 65).

Alderson uvádí zajímavé závěry výzkumu, jehož cílem bylo zjistit, zda výsledky testu při použití metody *cloze* ovlivňuje to, které slovo je z textu odstraněno jako první. „Výzkum ukázal, že pět různých verzí testu *cloze*, který byl vytvořen odstraněním každého pátého slova, a to počínaje slovem prvním, druhým, a tak dále vedlo k signifikantně odlišným testovým výsledkům“ (Alderson, 2000, s. 208).

Variantou této metody může být kombinace s výběrem odpovědi z nabídky, což lze realizovat v zásadě dvěma způsoby. Buď je pro každou mezeru nabídka několika možností výběru, nebo je pod testem uvedena banka možných odpovědí, jejichž počet by vždy měl převyšovat počet mezer.

Obr. 11 ukazuje testovací techniky společné a rozdílné pro ověřování úrovně poslechu a čtení. Z obrázku je zřetelné, že převážná většina technik je společná pro testování obou dovedností. Zapisování poznámek je jediná technika specifická pro testování poslechu, doplňování *cloze* zase jediná technika specifická pro testování čtení.



Obr. 11. Znárodnění společných a specifických testovacích technik pro zjišťování úrovně poslechu a čtení

## 5.2 Testování produktivních dovedností

Produktivní řečové dovednosti psaní a mluvení mají také mnoho společných rysů. Tato subkapitola se bude zabývat jejich definicemi, charakteristikou testování i typy testových úloh, které se v současné době nejčastěji používají k ověřování jejich úrovně.

### 5.2.1 Definice produktivních dovedností a jejich testů

Definice dovednosti psaní

V prostudované odborné literatuře nebylo možné najít žádnou definici dovednosti psaní. Na jedné straně se může zdát, že je zcela zjevné, co dovednost psaní představuje, a proto



není důležité snažit se o její definici. Naopak, při psaní zejména v cizím jazyce se může jednat o natolik složitý proces, že definovat ho může být příliš komplikované.

#### Definice dovednosti mluvení

Luoma uvádí, že „odborníci zabývající se vyučováním a testováním často hovoří o mluvení jako o termínu vztahujícímu se k různým dovednostem, které by učící se jazyku měli rozvíjet a mít“ (Luoma, 2004, s. 20). Fulcher definuje mluvení jako „verbální používání jazyka za účelem komunikace s ostatními“ (Fulcher, 2003, s. 23). Mluvení je jednou ze základních řečových dovedností a patří k dovednostem produktivním. Současné komunikativní přístupy k výuce cizích jazyků kladou na jeho rozvoj velký důraz.

#### Definice testu dovednosti psaní

Hughes se domnívá, že „nejlepší způsob, jak u jedinců testovat dovednost psaní je přimět je k tomu, aby něco napsali“ (Hughes, 1992, s. 75). Podle *Dictionary of Language Testing* je test psaní testem schopnosti psát v jazyce. Tato publikace uvádí dva hlavní přístupy k hodnocení psaní: nepřímé hodnocení, které vychází z objektivně hodnocených typů položek, kde je důraz kladen na pravopis, slovní zásobu a gramatiku, a přímé hodnocení, které má evidentně větší validitu než nepřímé testy. V tomto případě mají kandidáti napsat obvykle jeden až tři texty (Davies et al., 1999, s. 226-227).

V současné době je frekventovanější přístup druhý, což se jeví jako správné. Při hodnocení psaní je potřeba klást důraz na vypracování jasných kritérií hodnocení, aby byla zajištěna co nejvyšší objektivita hodnotitelů a reliabilita testových výsledků.

#### Definice testu dovednosti mluvení

Test dovednosti mluvení zjišťuje úroveň této dovednosti. Když chceme tuto úroveň hodnotit, snažíme se vést ústní projev kandidátů, a to prostřednictvím testových úloh, které vymezují obsah i kontext mluveného projevu (Luoma, 2004, s. 29). *Dictionary of Language Testing* vymezuje testování mluvení jako „hodnocení schopnosti mluvit v cílovém jazyce. Testy mluvení se pohybují v rozsahu od nepřímých .... k přímým“ (pozorování výkonu kandidáta v určitých interaktivních úlohách, jako například vysvětlování, stížnost, ...), (Davies et al., 1999, s. 182).

### 5.2.2 Charakteristika testování produktivních dovedností

#### Společné rysy testování produktivních dovedností

Na rozdíl od receptivních dovedností čtení i poslechu je produktivní dovednosti psaní a mluvení možné hodnotit přímo. Při hodnocení produktivních dovedností se však většinou nepoužívají objektivní testovací metody. Aby se maximálně zabránilo subjektivnosti hodnocení produktivních dovedností, je potřeba pečlivě připravit a ověřit v praxi hodnotící škály.

V souvislosti s testováním produktivních dovedností se často hovoří o hodnocení výkonu (*performance assessment*). Výkonový test charakterizuje Fitzpatrick a Morrison jako „test, ve kterém je určitá kritériální situace simulována v daleko větším rozsahu než u běžných testů typu tužka-papír“ (Fitzpatrick; Morrison, in McNamara, 1996, s. 10). Tito autoři také dále uvádějí, že hodnocení výkonu se může zabývat jak hodnocením *procesů* (v oblasti jazykového testování se jedná např. o hodnocení mluvení), tak *produktů* (hodnocení písemného projevu).

McNamara nabízí vysvětlení „silného“ a „slabého“ významu výrazu *test výkonu*. Při užití tohoto výrazu v „silném“ významu jde v praxi zejména o kritéria hodnocení výkonu. Testová úloha je v tomto případě reprezentována velmi realistickým, praktickým úkolem a výkon má být hodnocen primárně na základě toho, jestli úloha reálného života byla splněna. U tohoto typu hodnocení výkonu v jazykovém testování by však nešlo pouze o zjišťování jazykové úrovně, protože celkové hodnocení výkonu závisí na několika proměnných, přičemž úroveň jazyka je pouze jednou z nich. Vedle toho se v jazykovém testování častěji setkáváme s hodnocením výkonu v „slabém“ významu. V tomto případě jsou testové úlohy podobné úlohám v reálném životě, nebo je simulují. Avšak při hodnocení výkonu není v centru pozornosti hodnotitele pouze splnění úkolu. Účelem takovýchto testových úloh je přimět kandidáta k používání požadovaného jazyka v písemném (nebo ústním) projevu (McNamara, 1996, s. 44).

Kritéria hodnocení písemného i ústního projevu jsou velmi důležitá. Závisejí od účelu testu a mohou být lingvistická i nelingvistická. Měla by být srozumitelná hodnotitelům testu, jakož i kandidátům, kteří test vykonávají. „Pokud je testovou úlohou například napsání dopisu, lingvistická kritéria mohou zahrnovat plynulost, koherenci, kohéznost a gramatickou přesnost. Nelingvistická kritéria mohou zahrnovat splnění úkolu nebo vhodnost úpravy dopisu“ (Davies et al., 1999, s. 38).

Při hodnocení produktivních dovedností psaní a mluvení se používají dva hlavní typy hodnocení – analytické a globální (holistické). Analytické hodnocení (*analytic scoring, analytic assessment*) se provádí tak, že „každá důležitá dimenze výkonu je posuzována samostatně a z dílčích skóreů pro každou dimenzi je sestaven celkový skóre“ (Standardy pro pedagogické a psychologické testování, 2001, s. 173). Globální/holistické hodnocení (*global assessment, global scoring, impressionistic assessment, holistic scoring*) je možné charakterizovat jako „postup při skórování testu nebo testové položky, který vychází z posouzení celkového výkonu podle specifických kritérií“ (Standardy pro pedagogické a psychologické testování, 2001, s. 174).

Aby byly výsledky hodnocení písemného nebo ústního projevu spolehlivé, je nutné, aby byli hodnotitelé pravidelně proškolení a aby psaní nebo mluvení hodnotili minimálně dva na sobě nezávislí hodnotitelé. Harris se odvolává na výsledky výzkumu, při kterém byla kompozice hodnocena pouze jedním hodnotitelem. Koeficient reliability byl pouze 0.25 (Harris in Hughes, 1992, s. 86).

## Specifické rysy testování dovednosti psaní

U dovednosti psaní nesmíme zapomínat na skutečnost, že kandidáti často neumí správně psát ani v mateřském jazyce, a proto jim psaní v cizím jazyce může způsobovat velké problémy. Na první pohled by se dalo očekávat, že navržení testových úloh pro zjišťování úrovně dovednosti psaní je mnohem jednodušší než například navržení položek s vícenásobnou volbou odpovědi. Například při použití eseje jako testové úlohy by se mohlo zdát, že stačí navrhnout téma a nechat kandidáty, aby ho vypracovali. Skutečnost však není tak jednoduchá (Alderson; Clapham; Wall 1995, s. 59). Kromě tématu je potřeba stanovit časový limit, případně požadovaný rozsah písemné práce. Dále je nutné kandidátům jasně vymezit účel písemné práce, tedy zda se jedná o popis, vyprávění, prezentování vlastního názoru, shrnutí apod. a jakému čtenáři je písemný projev určen. Kandidáti by měli také vědět, co se bude na písemné práci hodnotit, za co se budou body přidělovat a případně i za co strhávat. Pro hodnotitele je pak nezbytné zpracovat jasná kritéria hodnocení, která je potřeba spolu s testovými úlohami před použitím pretestovat.

## Praktické otázky spojené s testováním dovednosti psaní

V současné době se zvyšuje důležitost schopnosti správně psát, a to nejen v cizím jazyce, ale i v jazyce mateřském. Různé testy aplikují při testování psaní různé přístupy. Na jedné straně snad nikdo nezpochybňuje fakt, že dovednost psaní je dovedností produktivní a tudíž výsledek testu psaní poskytuje relevantní informace. Na straně druhé určité existují i jazykové testy určené takovým populacím, které tuto dovednost v cílové situaci potřebovat nebudou. Většina obecně zaměřených standardizovaných testů anglického jazyka však úroveň písemného projevu ověřuje.

Při přípravě testu hodnotícího psaní je potřeba zohledňovat:

- věk kandidátů;
- reálnou schopnost psát na určitá témata, v určitém stylu nebo žánru v mateřském jazyce vyplývající např. z věku kandidátů;
- jejich znalost gramatiky, lexiky, syntaxe a stylistiky;
- jejich obecné znalosti;
- jejich potřebu písemného projevu v reálném životě.

Při tvorbě testu budou muset tvůrci testu rozhodnout o délce (trvání) testu a o počtu testových úloh. Kromě toho by hodnotící škála měla jasně a srozumitelně charakterizovat úroveň jazyka, která se od kandidátů dosahujících určitou jazykovou úroveň očekává. Také je potřeba uvést, za co se při hodnocení body udělují a za co se naopak strhávají. Velký význam má také vyškolení hodnotitelů, které musí být nedílnou součástí přípravy testu. Hodnotitelé musejí umět správně aplikovat hodnotící kritéria. Školení by mělo zajistit jednoznačný výklad kritérií a poskytnout dostatečný prostor pro praktické hodnocení různých ukázek písemného projevu.

Při přípravě testu psaní se doporučuje:

- zahrnout do testu více testových úloh, čímž se pravděpodobně lépe pokryje konstrukt;
- zjišťovat pouze úroveň dovednosti psaní a ničeho jiného, jako např. kreativity, inteligence nebo obecných znalostí;
- přesně vymezit, co se u testové úlohy od kandidátů požaduje, aby kandidáti ve svém písemném projevu nevybočovali z určeného rámce (Hughes, 1992, s. 75-86);
- jasně vyjádřit účel psaní; je důležité stanovit kontext testové úlohy a předpokládaného příjemce/čtenáře psaní;
- navrhnout testovou úlohu/testové úlohy, která/které bude/budou odpovídat případným situacím, ve kterých se kandidáti mohou ocitnout v reálném životě;
- tematické zaměření by mělo odpovídat věku, zájmům a úrovni vzdělání kandidátů;
- pokud jsou instrukce napsány v cílovém jazyce, je nutno používat jazyk takové obtížnosti, která je nižší než testovaná úroveň.

#### Specifické rysy testování dovednosti mluvení

Heaton upozorňuje na důležitost testování dovednosti mluvení i z důvodu pozitivního zpětného vlivu na vyučování angličtiny. „V jedné zemi byla ústní část ponechána jako součást maturitní zkoušky jednoduše proto, aby se v posledních dvou ročnících střední školy alespoň trochu mluvilo anglicky – a to přestože test samotný byl nespolehlivým měřicím nástrojem, protože ho zadával vysoký počet nekvalifikovaných examinátorů“ (Heaton, 1991, s. 89).

Teoretická literatura se zabývá například rozdělením testových úloh zjišťujících úroveň dovednosti mluvení na různé typy. „Brown a Yule rozlišují čtyři základní typy informačního rozhovoru: popis, pokyn, vyprávění příběhu a vyjádření/obhájení vlastního názoru“ (Brown; Yule 1983 in Luoma, 2004, s. 31). Bygate podává ještě přesnější rozdělení typů testových úloh pro dovednost mluvení:

#### „Fakticky orientovaný rozhovor

- popis
- vyprávění
- pokyn
- srovnání

#### Hodnotící rozhovor

- vysvětlení
- obhájení
- předpovídání
- rozhodnutí“

(Bygate, 1987 in Luoma, 2004, s. 32).

Při vymezení různých úrovní dovednosti mluvení se často uvádějí také komunikační funkce, které má kandidát na určité úrovni zvládnout. Společný evropský referenční rámec rozděluje tyto funkce na makro a mikro funkce. Makro funkce jsou části mluvené řeči sloužící stejnému účelu, jako například popisování, vyprávění, komentování nebo vysvětlování. Mikro funkce se naopak vztahují k jednotlivým aktivitám, které proběhnou v rámci jedné interakce; v SERR jsou rozděleny do šesti základních skupin:

- „ poskytnutí faktické informace a požádání o faktickou informaci, např. popis, hlášení, požádání;
- vyjádření postojů a zjištění postojů, např. souhlas/nesouhlas, znalost/neznalost, schopnost, dovolení;
- naléhání, např. návrh, žádost, varování;
- komunikace ve společenském styku, např. získání pozornosti, oslovení, pozdravení, představení se;
- strukturování promluvy, např. úvod, shrnutí, změna tématu, závěr;
- odstranění komunikačních překážek, např. projevení nepochopení, žádost o pomoc, parafrázování“ (SERR in Luoma, 2004, s. 34).

#### Praktické otázky spojené s testováním dovednosti mluvení

Testování dovednosti mluvení je činnost komplexní a náročná. Před samotným prováděním testu zahrnuje přípravu vhodných testových úloh a popisu jednotlivých úrovní (hodnotící škálu). Po pretestování testových úloh a hodnotícího nástroje pak probíhá samotné testování. V jeho průběhu může kandidát mluvit s jiným kandidátem, se zkoušejícím, nebo může své odpovědi nahrát na záznam a tudíž mluví s anonymním hlasem nahrávky. Testové úlohy může zadávat zkoušející osobně, nebo mohou být prezentovány na nahrávce v jazykové laboratoři. Hodnocení pak může vykonávat zkoušející nebo hodnotitel, a to buď bezprostředně po vykonání zkoušky nebo, pokud se zkouška nahrává, v určité době po vykonání zkoušky. Doporučuje se, aby zkoušející i hodnotitelé byli vyškoleni a poté pravidelně monitorováni, případně opětovně proškoleni. Jenom takovýmto způsobem je možné zajistit, že výsledky testu budou spolehlivé a subjektivnost hodnocení bude minimalizována.

Z odborné literatury lze vyvodit řadu doporučení vztahujících se k přípravě testových úloh, jejímu zadávání a hodnocení ústního projevu. Následující seznam určitě není vyčerpávající:

- Stimuly pro ústní projev by měly být navrženy tak, aby kandidáti prezentovali vhodný rozsah ústního projevu na požadované úrovni.
- Obtížnost formulace testových úloh by měla být nižší, než je testovaná úroveň.
- Tématické oblasti by neměly být stresující ani urážlivé.
- Testové úlohy by měly reflektovat potřeby reálného života cílové populace.
- Testové úlohy by neměly od kandidátů vyžadovat představivost a kreativitu, protože tyto vlastnosti se většinou v jazykovém testu netestují.
- Testové úlohy a témata by neměly být takové, které kandidátům působí obtíže v mateřském jazyce.
- Zkoušející by měli omezit počet otázek, na které lze odpovědět ano nebo ne.
- Zkoušející by se měli chovat tak, aby kandidáti byli co nejméně stresováni.

- Během zkoušky není vhodné kandidáty opravovat ani dělat si obsáhlé poznámky. Obě situace mohou kandidáty zbytečně stresovat.
- Zkoušející by měli mluvit málo.
- Zkoušející by měli mluvit srozumitelně a ptát se jasné otázky.
- Zkoušející by neměli kandidáty přerušovat a měli by jim dát čas přemýšlet.
- Výběr examinátorů by měl být velmi pečlivý – zkoušející by měli být sympatičtí a flexibilní s dobrou znalostí cílového jazyka (Hughes, 1992, s. 106).

### 5.2.3 Testovací techniky produktivních dovedností

Testovací techniky společné pro psaní a mluvení

Shrnutí (*Summary*)

Tato testovací metoda byla již uvedena jako testovací metoda společná pro zjišťování úrovně receptivních dovedností. Pro jejich hodnocení se také tato metoda používá nejčastěji. V poslední době se však někdy používá také k testování dovednosti psaní a mluvení.

Při použití tohoto typu testové úlohy pro zjišťování úrovně produktivních dovedností si kandidát přečte nebo poslechne text/nahrávku a poté je požádán, aby jeho/její obsah ústně shrnul. Úloha je autentická, avšak problémem zůstává hodnocení. Pokud by se udělovalo jedno skóre, nešlo by pouze o hodnocení psaní nebo mluvení, ale také o hodnocení čtení nebo poslechu. Je možné také udělovat skóre separátně pro poslech i mluvení. Zjišťovat míru porozumění podle následné ústní produkce však může být velmi složité.

Jednou z možností, jak zabránit tomuto problému, je použít ke shrnutí text v mateřském jazyce. Test v cílovém jazyce se může zdát na první pohled vhodnější, avšak způsobuje problémy při hodnocení, jak již bylo zmíněno. V případě použití textu v mateřském jazyce může zase docházet k tomu, že kandidáti mohou místo shrnutí sklouznout spíše do překladu. I když k tomuto nedojde, je dost pravděpodobné, že se bude více než u jiných metod projevovat interference mateřského jazyka.

Používání obrázků

Obrázky jsou v testování produktivních řečových dovedností oblíbeným a často používaným stimulem. Je možné použít obrázek jeden, přičemž se od kandidáta požaduje jeho popis. Pokud se použije série obrázků, která ukazuje příběh, je možné od kandidáta očekávat vyprávění nebo napsání příběhu. U kandidátů s vyšší jazykovou úrovní je možné použít dva obrázky, které jsou si podobné, odlišné nebo zcela protikladné a úkolem kandidátů je tyto obrázky srovnat.

Underhill shrnuje nejdůležitější výhody a nevýhody této techniky v ústním testování, které se však mohou projevit i při aplikaci této techniky v testování psaní. Za jednu z výhod považuje Underhill to, že stimulem pro mluvení není psaný text ale obrázek, a proto se nekladou současně požadavky na kandidátovu dovednost čtení; další výhodou je,

že všichni kandidáti reagují na stejný obrázkový stimul a tudíž jsou vzorky jejich řeči srovnatelné a je jednodušší je hodnotit. Kladně hodnotí také skutečnost, že slova, která kandidáti používají, nejsou úplně předurčena a existuje tady prostor k samostatnému vyjádření. Domnívá se, že dobrý obrázek dodává kandidátovi sebedůvěru a vyvolá „tok konverzace“. Poukazuje však i na některá úskalí používání této techniky. Může se stát, že kandidát nepochopí příběh znázorněný na obrázcích, a to buď z důvodů osobních nebo kulturních. Dále, pokud se kandidátům neposkytne klíčová slovní zásoba, může být kandidát, který zná alespoň několik klíčových slov z obrázku, ve velké výhodě vůči ostatním. Nakonec Underhill upozorňuje, že dobrý kreslený příběh pro tyto účely se hledá těžko (Underhill, 1994, s. 67). Variací použití obrázků v ústním testování je používání map a grafů jako stimulů k ústní produkci. Popis grafu může být následován delším ústním projevem, který nastiňuje možné další trendy. Tento typ stimulů je vhodný zejména pro testování ústního projevu pro akademické účely (Fulcher, 2003, s. 73).

#### Vyplnění dotazníku/formuláře (*filling in a questionnaire/form*)

Tato metoda se může použít jak pro zjišťování úrovně dovednosti psaní, tak i mluvení. Pro účely testu psaní je možné zkonstruovat dotazník nebo formulář, který kandidáti vyplní, případně lze použít autentický dotazník/formulář. Musíme však počítat s tím, že produkce písemného nebo ústního projevu může být kontaminována dovedností čtení.

Při použití této metody pro zjišťování úrovně dovednosti mluvení může probíhat vyplňování dotazníku/formuláře dvěma způsoby. Buď má dotazník/formulář k dispozici zkoušející a ptá se kandidáta otázky, aby spolu dotazník/formulář vyplnili, nebo má dotazník/formulář k dispozici kandidát a klade otázky zkoušejícímu. Pokud odpovědi získává a zapisuje kandidát, je potřebné počítat s kontaminací dovednosti mluvení s dovednostmi psaní a poslechu. Pozitivní vlastností této testové metody je, že je autentická. V reálném životě formuláře vyplňujeme často, proto tato činnost v cizím jazyce nepůsobí uměle a nepřírozně.

#### Informační transfer (*information transfer*)

Tento typ testové úlohy byl již zmíněn jako jedna z testovacích metod společných pro testování receptivních dovedností. Používá se taktéž v testování produktivních řečových dovedností. Při hodnocení úrovně dovednosti psaní a mluvení se tato metoda zaměřuje na přenos informací z grafické formy do formy ústního nebo písemného projevu. „Charakter kandidátových odpovědí pro tento druh produktivní úlohy je méně předvídatelný než u poslechu a čtení, což vede k vyššímu stupni subjektivity v hodnocení“ (Davies et al., 1999, s. 83). Lze říci, že tato testová úloha je autentická, avšak problémem může být spolehlivé a objektivní hodnocení.

#### Testovací techniky specifické pro testování dovednosti psaní

Do testu zjišťujícího úroveň dovednosti psaní se doporučuje zahrnout více než jednu testovou úlohu, nedávat kandidátům výběr ze dvou různých témat a hodnotit psaný projev minimálně dvěma hodnotiteli. Testovací metody nejsou natolik různorodé, jak je tomu

u receptivních dovedností. I když stanovit vhodnou testovací metodu a téma psaného projevu je bez sporu důležité, neméně důležité jsou hodnotící škály a jejich správné používání. Studenti by měli být informováni o tom, za co jim budou body přidělovány, resp. strhávány. Uvádíme pouze nejčastěji používané metody; v praxi jich existuje daleko širší škála. Jejich použití by vždy mělo reflektovat potřeby kandidátů.

Stimuly pro testové úlohy mohou obsahovat například sérii obrázků, pasáž k přečtení nebo citát. Výběr vhodného stimulu závisí od charakteristiky kandidátů. U kandidátů s nižší jazykovou úrovní pravděpodobně upřednostníme použití obrázků zobrazujících nějaký příběh před použitím například citátu nebo aforizmu – tento typ stimulu by byl mnohem vhodnější pro kandidáty s vyšší jazykovou úrovní.

#### Esej (*essay*)

Psaní esejů je v naší kultuře populární způsob zjišťování úrovně psaní jak v mateřském tak v cizím jazyce. Při navrhování takovéto testové úlohy je potřebné, aby kandidáti věděli, komu je esej určena. Další důležitá informace pro kandidáty je potřebný rozsah eseje, případně jestli se za nedostatečný nebo přebývající rozsah budou strhávat body.

#### Dopis (*letter*)

Psaní dopisu je velmi praktickou ukázkou psaní, kterou v reálném životě kandidáti často vykonávají. Je důležité specifikovat, jestli se jedná o dopis formální nebo neformální, jak dlouhý má být, jaké informace má obsahovat a komu je adresován.

#### Testovací techniky specifické pro testování dovednosti mluvení

Výběr testových úloh pro zjišťování dovednosti mluvení se řídí mnoha faktory. Testové úlohy, jejich obtížnost a tématické zaměření by měly reflektovat požadavky formulované v testových specifikacích a měly by tedy odpovídat účelu testu. Podobně jako u testování ostatních řečových dovedností je vhodné použít více než jeden typ testových úloh. Podle Hughese je „základní problém testování ústního projevu v principu stejný jako základní problém testování písemného projevu. Chceme vytvořit úlohy, které představují reprezentativní vzorek populace ústních úloh, o kterých se předpokládá, že je kandidát bude schopen splnit“ (Hughes, 1992, s. 101).

#### Interview, rozhovor (*interview*)

Tento typ testové úlohy se v ústním testování používá velmi často. Příprava a způsob provádění se však může lišit, a to od volného rozhovoru mezi kandidátem a examinátorem až k velmi strukturovanému typu interview. Je vhodné, aby interview mělo určitou strukturu a aby tedy bylo zajištěno, že je vedeno srovnatelným způsobem se všemi kandidáty. Na druhou stranu by však examinátorům měla být také poskytnuta jistá volnost při vedení rozhovoru, aby bylo dosaženo větší přirozenosti a aby kandidáti mohli mluvit volněji. Heaton varuje před přílišnou rigiditou struktury a uvádí příklad: „student



může rozvinout určité téma a postupovat v určitém směru, když ho zkoušející přeruší a zastaví celý tok rozhovoru, aby zahrnul otázku *Jak často*“ (Heaton, 1991, s. 101).

Rozhovor definuje *Dictionary of Language Testing* jako „termín označující přímé hodnocení úrovně řečových dovedností, přičemž každý kandidát je v interakci s jedním (nebo více) examinátory. Ústní interview je hodnoceno subjektivně, obvykle s použitím speciálních kritérií nebo s použitím holistické škály. Může se hodnotit okamžitě nebo pomocí audio nebo video záznamu výkonu“ (Davies et al., 1999, s. 136).

Ve světě známou ústní zkouškou je americká zkouška Oral Proficiency Interview (OPI), jejíž jednotlivé popisy úrovní vycházejí z popisů zpracovaných specialisty z Federal Service Institute (FSI) ve Spojených státech. Toto interview má několik fází. Začíná se fází „rozehřátí“ (*warm-up*), při které se zkoušející snaží navázat kontakt s kandidátem a zmírnit jeho nervozitu a napětí. Poté kandidát postupně dostává úlohy, jejichž pomocí examinátor zjišťuje úroveň kandidáta (*level checks*). V případě, že kandidát splňuje požadavky určité úrovně, examinátor mu zadává úkoly vyšší úrovně až po tzv. strop (*ceiling*). Poté co kandidát dosáhne svého stropu, může examinátor použít ještě sehraní situační role, aby se ujistil o správném stanovení úrovně kandidáta. Potom následuje závěr (*wind-down*), kdy examinátor interview uzavře pozitivním způsobem. Zajímavé při této zkoušce je, že její doba trvání je u různých kandidátů různá. Tak například u začátečníka může trvat 10 až 15 minut, zatímco u pokročilého kandidáta i 45 minut. Zkouška je nahrávána na audio kazetu. Kandidát je hodnocen nezávisle dvěma hodnotiteli. V případě, že se neshodnou, je kazeta předána hodnotiteli třetímu, který o výsledku definitivně rozhodne.

Zkouška vedená formou interview je velmi frekventovaná a má své výhody i nevýhody. Její obhájci se domnívají, že „nabízí realistický prostředek hodnocení celkové dovednosti mluvení v přirozené řečové situaci“ (Heaton, 1991, s. 97). Za nevýhodu je možné pokládat to, že je hodnocena subjektivně a tudíž je nutné dohlížet na reliabilitu výsledků. Heaton uvádí také další argumenty kritiků interview, kteří tvrdí, že kandidáti jsou u zkoušky v nepřírodných situacích, které je mohou psychologicky nepříznivě ovlivnit (Heaton, 1991, s. 97). Hughes upozorňuje na to, že při tomto typu ústní zkoušky testovaný vždy vystupuje v roli podřízeného a nemá tudíž prostor k projevení vlastní iniciativy. Proto jako možnou variantu tohoto typu testové úlohy uvádí interakci dvou nebo více kandidátů (Hughes, 1992, s.104-105). V tomto případě je sice pravdou, že kandidát nekomunikuje s examinátorem a tudíž není v pozici podřízené, mohou však vzniknout jiné problémy. Pokud examinátor do této diskuse nevstupuje, je možné, že dominantní kandidáti budou mluvit více než plašší kandidáti. Hodnocení pak může být těmito skutečnostmi ovlivněno.

Interview může být vedeno buď volnou formou nebo examinátoři mohou mít striktně předepsané otázky (skript), aby bylo hodnocení kandidátů jednodušší. Fulcher upozorňuje na to, že v takovémto případě kandidát víceméně odpovídá na otázky kladené examinátorem, což „limituje druh inferencí, které mohou být vyvozeny z počtu bodů“ (Fulcher, 2003, s. 79).

Odpovědi na nahrané otázky (*response to tape-recordings, tape-recorded stimuli, tape-based test, tape-mediated test*)

Tento způsob testování u nás není běžný, většinou ani není známý. Používá se zejména v případě, kdy je potřeba vyzkoušet vysoký počet kandidátů v krátké době a co nejefektivněji. Při použití tohoto typu testování je nezbytné věnovat velkou pozornost přípravě otázek, protože by měly stimulovat kandidáty k velmi podobným odpovědím. Takovéto testování probíhá v jazykové laboratoři, kde kandidáti odpovídají na stejné otázky z audio nahrávky. Tato technika by měla zajistit vysokou reliabilitu, protože na rozdíl od jiných typů testových úloh dovednosti mluvení je vysoce strukturovaná a prezentuje všem kandidátům naprosto stejné otázky. Jejich odpovědi jsou nahrávány a hodnoceny později. Underhill dále uvádí, že nahrané ústní testy jsou vhodné například v situacích, kdy může být organizačním problémem setkání kandidátů a examinátorů ve stejnou dobu na stejném místě, nebo v případě, že kandidáti jsou mladí lidé a mluvit s cizími lidmi by je mohlo uvádět do rozpaků (Underhill, 1994, s. 35). Nevýhodou tohoto typu testování je naopak skutečnost, že kandidáti reagují na anonymní hlas, je vyloučena spontánní interakce a testová situace není podobná situaci reálného života. *Dictionary of Language Testing* připouští výhodu tohoto typu testování ve standardizovaných podmínkách zadávání testu, čímž se vyloučí jakýkoliv vliv examinátora. „Na druhé straně vznikají otázky validity tohoto procesu, který nezahrnuje živou interakci“ (Davies et al., 1999, s. 195).

Prezentace (*presentation, oral report*)

Kandidáti si předem připraví prezentaci na určité téma. Tento typ testové úlohy se většinou používá při hodnocení ve specializovaných jazykových kurzech než ve velkých testovacích centrech. Přestože si kandidáti mohou své vystoupení připravit předem, jedná se pouze o přípravu stručných poznámek a čtení celých vět/textu se netoleruje. Důležité je správně si zvolit téma prezentace; musí být vybráno s ohledem na charakteristiku publika. V žádném případě se nejedná o předem naučené „povídání“ na určité téma.

Úloha zaměřená na získání informace (*information gathering task*)

Při použití tohoto typu testové úlohy se od kandidáta očekává iniciativa při zjišťování informací. Examinátor podá pouze stručnou informaci typu „Nedávno jsem byla v Řecku“ a úkolem kandidáta je zjistit co nejvíce informací týkajících se této skutečnosti. Kandidáti si mohou dělat poznámky, protože poté získané informace poskytnou druhému examinátorovi, a to buď v mateřském nebo cílovém jazyce. Úloha je autentická a vyžaduje od kandidáta iniciativu – tvoření otázek.

Úloha zaměřená na doplnění informace (*information gap task*)

Při tomto typu testových úloh kandidát obvykle dostane kartičku s informacemi, které má získat od examinátora, případně od jiného kandidáta. Tato aktivita je velmi autentická a kandidát není pouze v roli dotazovaného, jako například u rozhovoru (interview), ale sám musí otázky tvořit, aby informaci získal (podobně jako v úloze

zaměřené na získání informace). Fulcher tento typ úkolu hodnotí kladně a za hlavní výhodu pokládá skutečnost, že „kandidáti musejí 'opravdicky komunikovat'“ (Fulcher, 2003, s. 82).

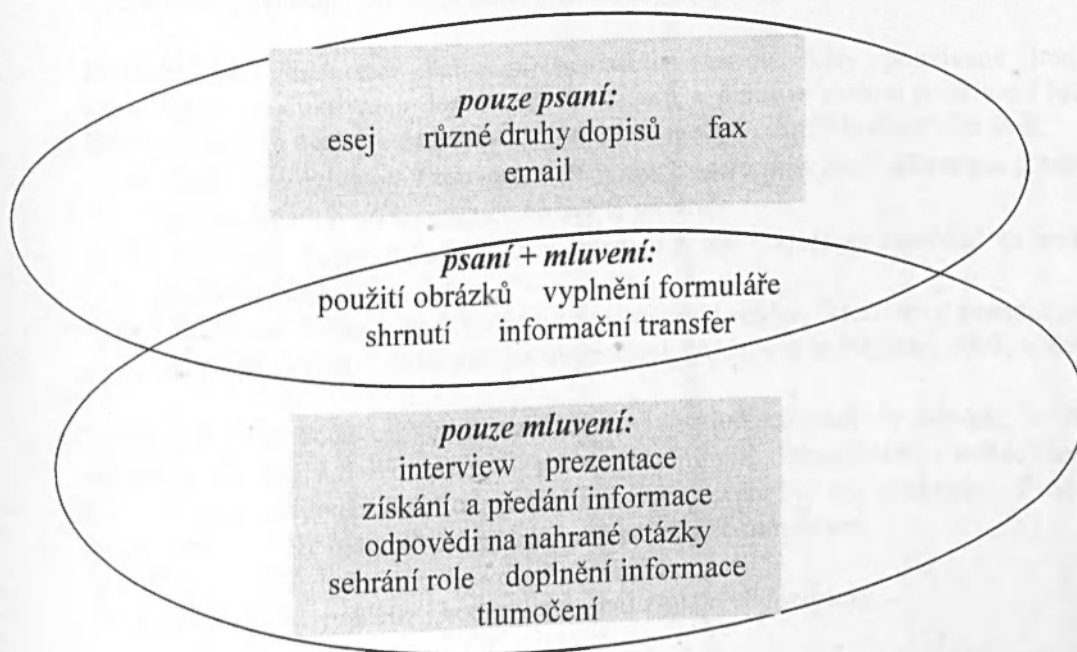
### Sehrání situační role (*role-play*)

U této testové úlohy kandidát sehrává určitou roli s examínátorem, případně s jiným kandidátem. Role je vymezena buď velmi přesně nebo volněji a může být různé obtížnosti – od nejjednodušších rolí až k rolím s různými komplikacemi. Kandidáti by měli být seznámeni jednak se samotnou testovou úlohou obecně a také s konkrétní situací a s rolí, kterou v ní mají sehrát. Sehrání situační role je úlohou autentickou. Při jejím návrhu je vždy potřeba brát v úvahu věk kandidátů, jejich znalosti a zájmy. Nemělo by se od nich vyžadovat sehrání takové role, kterou by nezvládli v mateřském jazyce.

### Tlumočení (*interpreting*)

Tento způsob testování ústního projevu není běžný. Kandidát tlumočí do nebo z cílového jazyka. Je možné, že pro určité specifické účely může mít tato technika své opodstatnění. Fulcher se však domnívá, že „v jazykovém testování je obvyklá snaha oddělit hodnocení „jiných dovedností“. Tento typ úlohy si zaslouží dalšího zkoumání, aby se mohl lépe definovat konstrukt, který testuje, což bude pravděpodobně směs konstruktů vztahujících se k mluvení, k poslechu a paměti“ (Fulcher, 2003, s. 84).

Obr.12 znázorňuje společné a odlišné testové úlohy pro psaní a mluvení. Na rozdíl od testovacích technik poslechu a čtení je společných typů testových úloh ověřujících psaní a mluvení jenom několik – pro srovnání viz také obr. 11.



Obr. 12. Znázornění společných a specifických testovacích technik pro zjišťování úrovně psaní a mluvení

#### 5.2.4 Hodnotící škály (*rating scales*)

Hodnotící škála (*rating scale, proficiency scale*) je podle *Dictionary of Language Testing* „škála určená k popisu dosažených jazykových znalostí a řečových dovedností obsahující sérii definovaných úrovní, vůči kterým se posuzuje výkon učícího se“ (Davies et al., 1999, s. 153). Písemný i ústní projev je nejčastěji hodnocen subjektivně. Při hodnocení písemného i ústního projevu je kromě výběru typů testových úloh a jejich návrhu také velmi důležité navržení hodnotících škál. Tyto škály obvykle popisují jednotlivé úrovně produktivních řečových dovedností od nulové úrovně až po úroveň vzdělaného rodilého mluvčího. Hughes však upozorňuje na to, že „užívání úrovně rodilého mluvčího, vůči které se posuzuje výkon (kandidátů), je v některých kruzích expertů na jazykové testování podrobován v poslední době kritice“ (Hughes, 1992, s. 110).

Hodnotící škály mohou obsahovat čísla, písmena, nebo stručné popisy úrovně (velmi dobrý, dostatečný apod.). Za základní typy hodnotících škál se pokládá škála holistická a analytická. Pokud používáme holistickou škálu (*holistic scale, impression scale*), jedná se o celkové hodnocení písemného nebo ústního projevu, přičemž se při hodnocení nesledují samostatně jednotlivé aspekty produkovaného projevu. Pokud se při hodnocení zaměřujeme na určité aspekty jazyka, které jsou separátně hodnoceny, jedná se o hodnotící škálu analytickou (*analytic scale*). Kromě těchto dvou základních hodnotících škál existuje v jazykovém testování mnoho jiných (Alderson; Clapham; Wall 1995, s. 107-108).

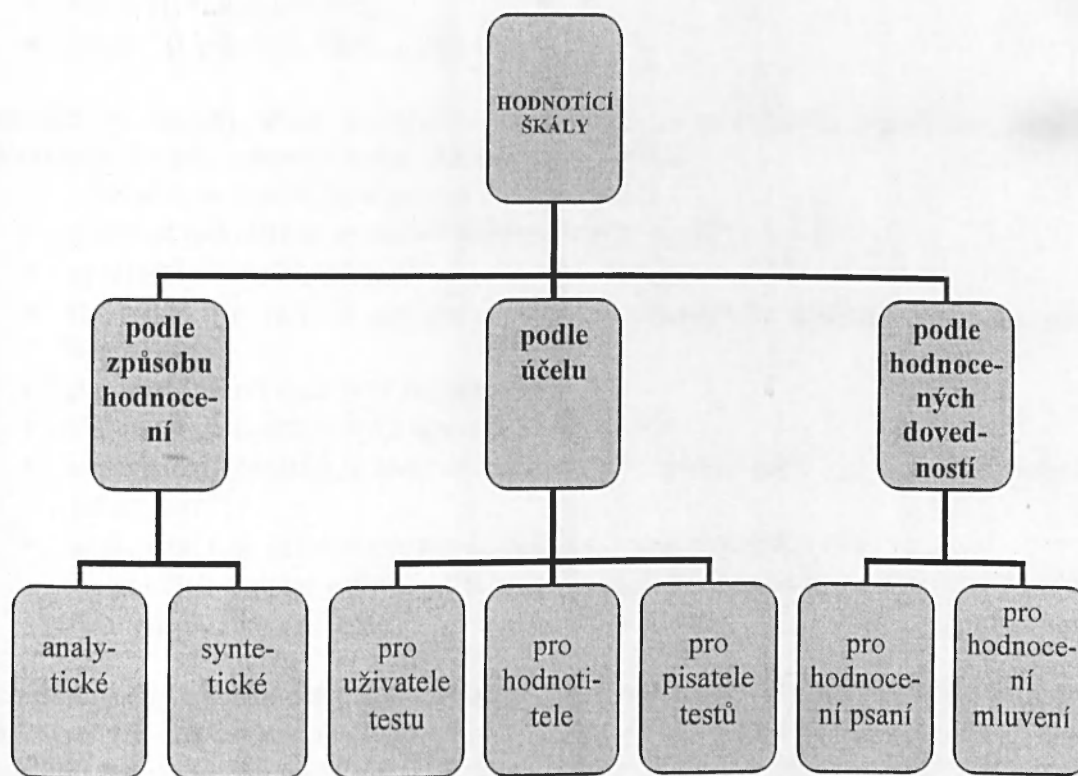
Obecně se nedoporučuje, aby hodnotící škály obsahovaly pouze čísla, případně jenom stručnou, jednoslovnou charakteristiku úrovně. Také se nedoporučuje používat více než sedmibodovou škálu. Dále je vhodné, aby pro každou konkrétní testovou úlohu byla vypracována zvláštní hodnotící škála (Alderson; Clapham; Wall 1995, s. 111).

Přestože jako hodnotící škály se nejčastěji chápou škály používané hodnotiteli k hodnocení produktivních dovedností kandidátů, v praxi se mohou používat i hodnotící škály s odlišným účelem. Alderson zmiňuje tato možná použití hodnotících škál:

- škály pro uživatele (*user-oriented scales*), které poskytují informace o tom, čeho je kandidát obvykle schopen na určité úrovni;
- škály pro hodnotitele (*assessor-oriented scales*), které se zaměřují na hodnocení výkonu kandidáta u testu;
- škály pro pisatele testů (*constructor-oriented scales*), které mají pomoci autorům testu při výběru vhodných testových úloh (Alderson in Fulcher, 2003, s. 89).

Pokud má počet bodů spolehlivě odrážet úroveň produktivních dovedností, je důležité, aby hodnotící škály byly napsány jasně a srozumitelně. Examinátoři i hodnotitelé musí být pro správné používání škál proškoleni jak teoreticky, tak prakticky. Produktivní dovednosti by měly být hodnoceny více než jedním hodnotitelem.

Obr. 13 znázorňuje rozdělení hodnotících škál z různých hledisek.



Obr. 13. Rozdělení hodnotících škál podle různých hledisek

Každý testovací systém má svoji vlastní hodnotící škálu, která je v souladu s účelem testu a odpovídá testovým specifikacím. Hodnotící škála se soustřeďuje na ty aspekty jazyka, které jsou pro určitý test a určitou testovou úlohu relevantní.

V písemném projevu se nejčastěji hodnotí:

- gramatika,
- slovní zásoba,
- pravopis,
- plynulost projevu,
- organizace projevu,
- splnění komunikačního záměru,
- přesnost projevu,
- vhodnost projevu,
- relevance obsahu projevu,
- koheze.

V ústním projevu za nejčastěji hodnocené lze pokládat tyto oblasti:

- „gramatika,
- slovní zásoba,
- výslovnost, intonace a přízvuk,

- styl a plynulost projevu,
- obsah“ (Underhill, 1994, s. 96).

Specifičtější aspekty, které souvisejí se skutečností, že se u jazyka dává důraz na jeho charakteristiku jako nástroje komunikace, jsou například:

- „rozsah (jak dlouhé jsou pronesené výpovědi?)
- složitost (jak dalece se mluvčí snaží o složitý jazyk?)
- rychlost (jak rychle mluví?)
- flexibilita (je mluvčí schopen rychle se přizpůsobit změnám v tématu nebo úloze?)
- přesnost (mluví správnou angličtinou?)
- vhodnost (je funkční styl odpovídající/vhodný?)
- nezávislost (spoléhá mluvčí na otázku nebo stimul anebo je schopen iniciovat projev sám?)
- opakování (jak často je potřeba otázku nebo stimul zopakovat?)
- váhání (jak mnoho mluvčí váhá předtím než začne mluvit a během mluvení?)“ (Underhill, 1994, s. 96).

Hodnotící škály by také měly uvést, jaké chyby jsou v ústním projevu pro určitou úroveň přijatelné. Heaton se domnívá, že přísně by se měly penalizovat takové chyby, které narušují úspěšnou komunikaci. Mírnější postoj je podle něho potřeba zaujmout k chybám, které jsou méně závažné a komunikaci nenarušují (Heaton, 1991, s. 95). S tímto názorem je možné souhlasit při hodnocení ústního projevu na nižších úrovních. U vyšších úrovní je také vhodné určité chyby tolerovat; je však potřeba tyto chyby přesně definovat.

## 6. Testové úlohy z pohledu Bloomovy taxonomie

První kapitola této práce pojednává o vzdělávacích cílech a uvádí hlavní myšlenky, které stály za zrodem Bloomovy taxonomie. Poté stručně shrnuje další vývoj, který vedl k jejímu rozšíření. V této kapitole bude Bloomova taxonomie zmíněna konkrétněji a některé typy testových úloh popsané v předchozí kapitole budou z jejího pohledu charakterizovány. Jinými slovy, pokusíme se některé typy testových úloh používaných v jazykových testech zařadit do obecnějšího rámce, který Bloomova taxonomie poskytuje. Na rozdíl od první verze taxonomie, která byla jednorozměrná, je nová taxonomie dvourozměrná – kromě šesti kategorií kognitivních procesů obsahuje také čtyři kategorie znalostí. Navíc poskytuje čtenářům ukázky aplikace Bloomovy taxonomie v praxi.

Na horizontální ose se nacházejí kognitivní procesy, které jsou vyjádřeny šesti klíčovými slovesy: 1. zapamatovat, 2. porozumět, 3. aplikovat, 4. analyzovat, 5. hodnotit a 6. vytvořit. Vertikální osa je dimenzí znalostí, které jsou rozděleny na čtyři kategorie: A. faktické znalosti, B. znalosti pojmů, C. znalosti procesů a D. meta kognitivní znalosti. Tato kapitola se bude zabývat pouze kognitivními procesy a testovacími technikami,

kteřé je mohou v jazykovém testování ověřovat. V této souvislosti nebude snad nezajímavé zmínit se o některých myšlenkách Muellera, jehož definice vzdělávacích standardů a cílů byly zmíněny v kapitole 3. Mueller se kromě toho zabývá rozdílem mezi tradičními a autentickými testovými úlohami (viz subkapitolu 3.2), a právě tyto rozdíly budou uvedeny ve srovnání s Bloomovou taxonomií.

Kategorie **zapamatovat** (*remember*) představuje „znovunalezení relevantních znalostí v dlouhodobé paměti“ (Anderson; Krathwohl 2001, s. 66). Jako dva kognitivní procesy spojené s touto kategorií jsou uváděny procesy **rozpoznání** (*recognizing*) a **vybavení si** (*recalling*). Tato kategorie je uváděna jako první, základní, tedy nejjednodušší. Odpovídajícími testovacími technikami, které zjišťují úroveň tohoto kognitivního procesu při jazykovém testování, by mohly být: technika výběru správné odpovědi z vícenásobné volby, pravdivá a nepravdivá tvrzení nebo různé druhy přiřazování. Tyto typy testových úloh klasifikuje Mueller jako testové úlohy s vybranou odpovědí, ve kterých se netvoří nic nového a studenti si vybaví nebo rozpoznají správnou odpověď. V jazykovém testování jsou to techniky, u kterých kandidáti jazyk neprodukují v písemné ani v ústní podobě, pouze vybírají odpovědi z hotové nabídky. Tuto kategorii lze považovat za relevantní u metody tzv. *discrete points testing*, kdy se overuje zvládnutí např. jednotlivých lexikálních jednotek nebo gramatických jevů. K této kategorii lze přiřadit atributy tradičního hodnocení podle Muellera – výběr odpovědi jako opozitum k vykonání (provedení) úlohy, vymyšlená úloha jako protiklad úlohy z každodenního života, rozpoznání a vybavení si jako protiklad ke konstrukci a aplikaci vědomostí (<http://jonathan.mueller.faculty.noctrl.edu/toolbox/tasks.htm> 10.8.2006). Přestože Mueller (<http://jonathan.mueller.faculty.noctrl.edu/toolbox/tasks.htm> 10.8.2006), nezavrhuje ani testování, které obhájuje autentické hodnocení (*authentic assessment*), domnívá se, že testy podle Bloomovy taxonomie ověřuje kategorii zapamatovat. Mueller se domnívá, že testy zjišťující zapamatování nebo vybavení si jsou vhodným doplněním autentického testování a také připouští i jistou autentičnost testových úloh, ve kterých je potřeba rozpoznat nebo vybavit si určitá fakta, protože i to se od nás v reálném životě očekává (<http://jonathan.mueller.faculty.noctrl.edu/toolbox/tasks.htm> 10.8.2006).

Testové úlohy ověřující další, v této taxonomii vyšší kognitivní procesy, musí být konstruovány takovým způsobem, aby nebylo možné na ně správně odpovědět pouze na základě zapamatovaného (Anderson; Krathwohl 2001, s. 71).

Druhou kategorií nové Bloomovy taxonomie je **porozumět** (*understand*). „Studenti rozumí tehdy, když si vytvářejí spojení mezi „novými“ znalostmi, které mají získat a předchozími znalostmi“ (Anderson; Krathwohl 2001, s. 70). S touto kategorií se pojí kognitivní procesy interpretace, uvedení příkladu, rozřídění, shrnutí, odvozování, srovnání a vysvětlení. Proces **interpretace** (*interpreting*) by v jazykovém testování bylo možné ověřit například technikou výběru správné odpovědi z vícenásobné volby, technikou pravdivých a nepravdivých tvrzení, technikou přiřazování, informačního transferu (přenos slyšených nebo přečtených informací do tabulky, grafu apod.) nebo technikou *cloze*. Jednalo by se o testování receptivních dovedností.

**Uvedení příkladu** (*exemplifying*) a **rozřídění** (*classifying*) je možné ověřovat v jazykovém testování technikou výběru správné odpovědi z vícenásobné volby,

technikou pravdivých a nepravdivých tvrzení a krátkých odpovědí. Jedná se o techniky používané k testování receptivních řečových dovedností.

Naopak *shrnutí* (*summary*) je možné v jazykovém testování ověřit technikou výběru správné odpovědi z vícenásobné volby nebo doplněním chybějících informací v textu. Tyto techniky je možné aplikovat v hodnocení receptivních dovedností.

Pro *odvozování* (*inferring*) je možné najít ekvivalent v doplňovacích testovacích technikách, například v technice doplňování mezer, v technice *cloze* a také v technice shrnutí *cloze*. Navíc, pokud jsou správně zkonstruovány, mohou tuto kategorii zjišťovat také testovací techniky s uzavřenou odpovědí (výběr odpovědi z vícenásobné volby, pravdivá a nepravdivá tvrzení, otázky s krátkou odpovědí). Kromě toho je možné vidět také paralelu mezi tímto kognitivním procesem a dílčí řečovou dovedností – odvozování významu neznámých slov.

*Srovnání* (*comparing*) a *vysvětlení* (*explaining*) nejsou jako takové předmětem ověřování běžnými testovacími technikami jazykového testování. Mohou se spíše používat jako požadované jazykové funkce.

Třetí kategorie je nazvána **aplikovat** (*apply*) a jako další kognitivní procesy s ní spojené jsou uvedeny *provádění* (*executing*) a *zavádění* (*implementing*). Jako obecný příklad pro *provádění* je možné zmínit cvičení. „Cvičení je úloha, pro jejíž vykonání student zná správný postup.“ Takže použitý postup vykonává víceméně rutinně (Anderson; Krathwohl 2001, s. 77). Příkladem může být sehrání situační role, pokud se nepoužívá jako testovací technika, ale pouze jako cvičení. Podobně bychom mohli vidět paralely v cvičeních zaměřených na jazykové znalosti, například transformace vět kladných na záporné, oznamovacích na tázací, činných na trpné a podobně. Při testování jazyka je například u zkoušek vytvořených UCLES (University of Cambridge Local Examinations Syndicate) používaná technika tzv. transformace, kdy se věta má přeformulovat při zachování významu a jsou dány např. první dvě slova takové transformované věty.

Testovací technikou vhodnou pro *zavádění* v oblasti jazykového testování by mohla být například technika sehrání situační role, a to za předpokladu, že kandidáti sice tuto techniku znají, avšak neznají konkrétní zadání.

Následující kategorií je kategorie **analyzovat** (*analyze*), se kterou souvisejí procesy *diferenciace* (*differentiating*), *organizace* (*organizing*) a *přisuzování* (*attributing*). Při procesu diferenciaci „student odlišuje relevantní informace od irelevantních, nebo důležité od nedůležitých...“ (Anderson; Krathwohl 2001, s. 80). Tento proces má v jazykovém testování obdobu v dílčích receptivních dovednostech. Hodnotící nástroj má většinou formu uzavřených odpovědí (výběr odpovědi z vícenásobné nabídky, pravdivá a nepravdivá tvrzení, přiřazování). Konkrétně je možné zjišťovat, zda kandidáti identifikují hlavní myšlenku slyšeného nebo čteného textu, detailní informace, postoj mluvčího nebo pisatele apod. Proces organizace obvykle probíhá společně s procesem diferenciaci a student si v něm vytváří spojení mezi prvky prezentovaných informací (Anderson, Krathwohl 2001, s. 81). V jazykovém testování je tento proces možné měřit testovými



úlohami s uzavřenou odpovědí (výběr odpovědi z vícenásobné nabídky, pravdivá a nepravdivá tvrzení, přiřazování) a jedná se zejména o testování receptivních řečových dovedností. Další možnou technikou testování může být technika zapisování poznámek.

Jinou vhodnou technikou pro testování toho, jestli je kandidát schopen sledovat organizaci mluveného nebo psaného projevu, je technika seřazování vět nebo odstavců do uceleného textu (*jumbled text, ordering, rearrangement*).

Co se týká procesu přisuzování, většinou jde o zjištění postojů nebo záměru autora. Takovéto zjišťování je možné v jazykovém testování zejména u testů receptivních dovedností a jako testové úlohy jsou vhodné úlohy s uzavřenou odpovědí (výběr odpovědi z vícenásobné nabídky, pravdivá a nepravdivá tvrzení, přiřazování), případně úlohy s krátkou odpovědí.

Právě dvě předchozí kategorie (aplikovat, analyzovat) uvádí Mueller jako atributy autentického hodnocení: „...autentické hodnocení často požaduje od studentů, aby analyzovali, syntetizovali a aplikovali to, co se naučili“ (<http://jonathan.mueller.faculty.noctrl.edu/toolbox/tasks.htm> 10.8.2006).

Kategorie **hodnotit** (*evaluate*) je chápána jako „vytvoření úsudku na základě kritérií a standardů“ (Anderson; Krathwohl 2001, s. 83). Patří k ní proces **ověřování** (*checking*) a **posouzení** (*critique*). Proces ověřování je možné identifikovat v testování receptivních dovedností, a to například při použití testovací techniky pravdivých a nepravdivých tvrzení nebo použití výběru odpovědí z vícenásobné volby. K procesu posouzení není možné najít odpovídající ekvivalent v testovacích technikách.

**Vytvořit** (*create*) je nejvyšší kategorií v této taxonomii. Od studentů se očekává vytvoření, vyprodukování určitého celku. Přestože pro někoho může toto slovo vyvolávat představu tvorby něčeho neobvyklého, vztahuje se tato kategorie „k cílům, které požadují takovou produkci, kterou všichni studenti dokáží udělat a udělají“ (Anderson; Krathwohl 2001, s. 85). Kategorie **vytvořit** jinými slovy požaduje od studentů, aby vytvořili syntézu různých částí (například informací nebo materiálů), při které vznikne jednotný celek (Anderson; Krathwohl 2001, s. 85). Za tři hlavní části, vztahující se k produkci, se pokládá **generování** (*generating*), **plánování** (*planning*) a **produkování** (*producing*). Co se týká generování a plánování, v jazykovém testování nepředstavují samostatný předmět ověřování. Avšak proces produkování se očekává u převážné většiny jazykového testování ověřujícího úroveň produktivních řečových dovedností. Odpovídajícími testovými úlohami jsou např. shrnutí, ústní či písemný popis obrázků, informační transfer, napsání dopisu, napsání eseje, ústní rozhovor, ústní prezentace, sehrání situační role a další.

Mueller jako protiklad testových úloh s vybranou odpovědí uvádí testové úlohy s tvořenou/konstruovanou (*constructed*) odpovědí. Podle něho lze tyto testové úlohy rozdělit na produktové (*product-like*), například eseje, povídky, básně, výzkumné zprávy, postery atd. a výkonové (*performance-like*), například provedení experimentu, debata, ústní prezentace apod.

Přehled uvedeného pokusu o zařazení některých typů testových úloh do kategorií kognitivních procesů Bloomovy taxonomie ilustruje následující tabulka.

Tab. 3. Zařazení některých typů testových úloh do kategorií kognitivních procesů Bloomovy taxonomie

zapamatovat	porozumět	aplikovat	analyzovat	hodnotit	vytvořit
výběr odpovědi z vícenásobné volby	výběr odpovědi z vícenásobné volby	sehrání situační role	výběr odpovědi z vícenásobné volby	výběr odpovědi z vícenásobné volby	shrnutí
pravdivá a nepravdivá tvrzení	pravdivá a nepravdivá tvrzení	transformace	pravdivá a nepravdivá tvrzení	pravdivá a nepravdivá tvrzení	používání obrázků pro popis, srovnání apod.
přiřazování	přiřazování		přiřazování		informační transfer
	informační transfer		zapisování poznámek		napsání dopisu
	doplňování a cloze		úlohy s krátkou odpovědí		napsání eseje
	úlohy s krátkou odpovědí		seřazování		ústní rozhovor
					ústní prezentace
					sehrání situační role

#### Komentář

V tabulce se nejčastěji vyskytují techniky výběru odpovědi z vícenásobné volby (4 x) a pravdivých a nepravdivých tvrzení (4 x). Byly uvedeny u kognitivních procesů zapamatovat, porozumět, analyzovat a hodnotit. Z hlediska jazykového testování se jedná zejména o hodnocení dovedností receptivních.

Technika přiřazování se vyskytuje v tabulce třikrát (kognitivní procesy zapamatovat, porozumět a analyzovat), technika informačního transferu se vyskytuje dvakrát (kognitivní procesy porozumět a vytvořit). I tyto techniky se používají v hodnocení receptivních dovedností.

Ostatní typy testovacích technik se vyskytují pouze jednou. Kognitivnímu procesu aplikovat je přiřazen pouze jeden typ testovací techniky, a to sehrání situační role. Nejvíce technik je naopak přiřazeno kognitivnímu procesu vytvořit a jedná se pouze o techniky testující úroveň dovedností produktivních.

#### Dílčí závěr

Zařazení různých technik používaných v jazykovém testování do obecnějšího rámce Bloomovy taxonomie mělo za cíl zjistit, které z kognitivních procesů kandidáti při testování různými technikami s velkou pravděpodobností používají. Prezentovaná tabulka je originální, ale reprezentuje subjektivní názor, protože prostor k širší diskusi s učiteli a specialisty na testování v rámci této práce není. Přesto si myslím, že z uvedené tabulky lze obecně říci, že při testování receptivních řečových dovedností (poslech s porozuměním, čtení s porozuměním) kandidáti potřebují vykonávat jeden nebo více kognitivních procesů (*zapamatovat, porozumět, aplikovat, analyzovat, hodnotit*), avšak ne kognitivní proces *vytvořit*. Tento je potřebný pouze k vypracování nejrůznějších testových úloh ověřujících dovednosti produktivní (mluvení, psaní). Protože kognitivní proces *vytvořit* je v Bloomově taxonomii hierarchicky nejvyšší, domnívám se, že by testy vysoké důležitosti měly být vytvořeny tak, aby od kandidátů jeho používání vyžadovaly. U jazykových testů to znamená, že by kromě úrovně receptivních dovedností měly také zjišťovat úroveň dovedností produktivních. Proto jsem přesvědčena, že maturitní zkouška z cizího jazyka by měla testovat všechny čtyři základní řečové dovednosti (poslech s porozuměním, čtení s porozuměním, psaní a mluvení).

## II. NÁVRH ALTERNATIVNÍHO KATALOGU POŽADAVKŮ; KONSTRUKCE, ZADÁNÍ A VYHODNOCENÍ TESTU

### 7. Hodnocení Katalogu požadavků: pozitivní a negativní stránky

#### 7.1 Seznámení se s problematikou – zjištění získaná v dosavadní práci

Vzhledem k tomu, že nepracuji ve středoškolském prostředí, připravila jsem si v roce 2002 projekt, jehož cílem bylo zjistit, jaké jsou charakteristiky výuky anglického jazyka u maturitních ročníků středních škol, jaké metody výuky se používají ve třídách, jaká je forma a průběh maturitní zkoušky, jak je maturita z anglického jazyka vedena a jak a co se u maturitní zkoušky hodnotí. Do průzkumu byly zapojeny tyto brněnské střední školy: Gymnázium tř. kpt. Jaroše (GKJ), Gymnázium Matyáše Lercha (GML), Gymnázium Tábořská (GTB), Střední průmyslová škola chemická (SPŠCH), Střední průmyslová škola textilní (SPŠTX) a Integrovaná střední škola – centrum odborné přípravy (ISS-COP).

Průzkum se skládal z několika částí: 1. pozorování v hodinách anglického jazyka, 2. vypracování, zadání a vyhodnocení srovnávacího testu, 3. forma a obsah maturitní zkoušky a okruhy maturitních otázek a 4. pozorování maturity (viz příloha P - 2).

Jak ukázal průzkum popsáný v příloze P - 2, maturitní zkouška na pozorovaných školách nebyla vedena správně. Jediným materiálem, který dává informace jak učitelům tak žákům, jsou více-méně názvy tématických okruhů, označované jako maturitní otázky, případně texty, obvykle okopírované z učebnic angličtiny, jejichž porozumění se zjišťuje. Neexistuje dokument, který by odpověděl na zcela elementární otázky, jako například co zkouška testuje, jakými metodami to testuje a jaký výkon je považován za dostatečný ke složení zkoušky. Odpovědi na tyto otázky by měly být jasné jak učitelům, tak i žákům. Právě odpověďmi na tyto otázky se bude zabývat alternativní model *Katalogu požadavků*.

#### 7.2 Jaké jsou pozitivní vlastnosti Katalogu požadavků?

Katalog požadavků podává do určité míry ucelený a systematický přehled toho, co má maturitní zkouška za cíl ověřovat. Jeho obsah tvoří požadavky na cílovou jazykovou úroveň žáků, kteří ukončí své středoškolské vzdělání v anglickém jazyce maturitou. Je to první dokument, který tímto způsobem specifikuje požadavky na jazykové znalosti a řečové dovednosti. Tuto skutečnost pokládám za velmi důležitou, protože v současné době je vedení maturitní zkoušky z jazyka plně v kompetenci jednotlivých středních škol, což jim umožňuje stanovování velmi odlišných požadavků, jakož i kritérií hodnocení. Pak je jenom stěží možné, aby hodnocení maturitní zkoušky stejnou známkou z různých škol bylo srovnatelné. Vznik tohoto Katalogu požadavků lze označit jako velmi pozitivní skutečnost.

Kromě stanovení cílových požadavků na jazykovou úroveň maturantů Katalog požadavků seznamuje své čtenáře také s charakteristikou testu pro společnou část maturitní zkoušky, která uvádí délku testu, počet subtestů, typy použitých testových úloh i jazyk, ve kterém jsou zadávány instrukce k jednotlivým testovým úlohám. Typy testových úloh jsou většinou uzavřené, pouze malá část úloh může zahrnovat otevřené úlohy s krátkou odpovědí. Výhodou tohoto rozhodnutí je určitě skutečnost, že převážně objektivní test natolik neohrozí reliabilitu testových výsledků. Na druhou stranu uvedené typy testových úloh neumožňují žákům produkci jazyka, která je pro hodnocení jejich celkové jazykové úrovně velmi důležitá. Je však potřeba uvědomit si, že tyto testové úlohy se mají podle Katalogu požadavků používat k hodnocení receptivních řečových dovedností, a tedy danému účelu odpovídají (viz obr.10 a 11).

Kladem tohoto dokumentu je nepochybně také uvedení vzorových testů zjišťujících úroveň dovednosti poslechu a čtení se správným řešením. Jakkoliv přesný je popis požadavků na úroveň v jednotlivých řečových dovednostech, seznámení se s ukázkovými testovými úlohami je velmi důležité jak pro učitele, tak i studenty.

Katalog požadavků uvádí, že pro společnou část maturitní zkoušky bude možné volit ze dvou úrovní obtížnosti, a to základní a vyšší. Toto rozhodnutí má určitě své odůvodnění; jedním z nich je fakt, že za těchto podmínek nebudou znevýhodněni například žáci středních odborných škol – pokud totiž výuce jazyka byla věnována nižší hodinová dotace, budou se moci rozhodnout pro základní úroveň, jejíž obtížnost je nižší.

Obecně je možné pokládat za správné rozdělení řečových dovedností na receptivní a produktivní. V části CÍLOVÉ KOMPETENCE i SPECIFICKÉ CÍLE se autoři textu Katalogu snaží jasně a srozumitelně formulovat, co má žák umět v jednotlivých řečových dovednostech. Přesto se všemi požadavky na řečové dovednosti není možné souhlasit (uvedeno dále v subkapitole 7.3).

Odborná literatura z oblasti teorie testování zdůrazňuje význam publikování informačních brožur pro kandidáty. Doporučuje se, aby obsahovaly například informace o účelu testu, testových položkách a typech testových úloh, dále o formátu testu, obtížnosti a hodnotících kritériích (Davies et al., 1999, s. 207). Cílem zveřejnění těchto informací je, aby se studenti co nejvíce seznámili s testem předtím, než ho budou psát, což jim umožní podat co nejlepší výkony (Alderson; Clapham; Wall 1995, s. 21). Většinu těchto údajů Katalog požadavků obsahuje.

Pro srovnání je zajímavé uvést, že například u cambridgeských zkoušek je možné všechny důležité informace získat na internetu, a to včetně ukázkových testů (<http://www.cambridgeesol.org/support/handbooks.htm>). Kromě toho jsou samozřejmě komerčně dostupné jak výukové materiály, tak cvičné testy. V této souvislosti určitě také stojí za zmínku, že nové Katalogy požadavků, které obsahují i ukázky testových úloh pro jednotlivé řečové dovednosti, jsou v současnosti již také dostupné na internetu (<http://www.ceremat.cz/katalogy/>, <http://www.edu.cz>, <http://www.msmt.cz>).

### 7.3 Které charakteristiky Katalogu požadavků je potřeba zdokonalit?

Tato část se zabývá hlavními nedostatky Katalogu požadavků (dále Katalog). Přestože je uvedeno víc nedostatků, nelze zapomínat, že tento dokument byl první svého druhu a tudíž je přirozené, že nemohl být ideální. Navíc nový Katalog je již na daleko vyšší formální i obsahové úrovni.

#### 7.3.1 Obecné formulace

Předmluva ke Katalogu uvádí, že tento dokument je určen žákům středních škol, jejich rodičům, učitelům středních škol, institucím terciárního vzdělávání, zaměstnavatelům, široké veřejnosti i zahraničním partnerům. Obávám se však, že terminologie a formulace používané v Katalogu nejsou všem těmto skupinám srozumitelné (např. komunikativní kompetence; společensko-politické reálie; komunikační záměr; kompetence jazyková, pragmatická, strategická, sociolingvistická; kompenzační vyjadřování apod.).

Z Katalogu není jasné, jak bude probíhat samotné hodnocení maturitní zkoušky, kdo ho bude provádět, jaký podíl na známce budou mít počty bodů za výkony v jednotlivých řečových dovednostech, jestli například hodnocení úrovně znalostí jazyka (Katalog uvádí pod termínem jazyková kompetence) bude mít stejnou váhu jako hodnocení úrovně dovednosti mluvení. Není jasné, jestli výsledky maturitní zkoušky budou vyjádřeny jednou celkovou známkou, několika známkami z jednotlivých subtestů, jazykovým profilem s vysvětlením anebo nějakým jiným způsobem. Informace o způsobu hodnocení, případně o vážení výsledků u jednotlivých subtestů lze pokládat za velmi důležité a v tomto dokumentu zjevně chybějí.

Přestože Katalog uvádí, že obě části maturitní zkoušky, tedy společná a školní, jsou rovnocenné, školní část maturity je víceméně jenom zmíněna a není popsána stejně podrobně jako část společná. Z Katalogu se dozvíme, že školní část maturitní zkoušky se bude zabývat zejména hodnocením produktivních řečových dovedností, tedy mluvení a psaní. Kromě toho se od školní části maturity očekává, že bude hodnotit nejenom řečové dovednosti a jazykové prostředky, ale také předpokládá zvládnutí „dalších naukových složek předmětu (reálií, kulturního přehledu, podle zaměření školy literatury nebo odborné terminologie)“ (Katalog požadavků, 2000, s. 5). Z tohoto důvodu by bylo vhodné také blíže vymežit, co z dalších zmíněných naukových složek předmětu se bude hodnotit a jakým způsobem. Pro srovnání stojí za zmínku, že testování reálií v rámci mezinárodně certifikovaných jazykových zkoušek není obvyklé, protože hodnocení se pak stává velmi složitým a nepřehledným. Není totiž jasné, jestli se hodnotí řečové dovednosti nebo znalosti reálií.

#### 7.3.2 Receptivní dovednosti

Řečové dovednosti jsou v Katalogu rozděleny na receptivní, produktivní a mediační. Nejvíce prostoru věnuje Katalog popisu požadavků na receptivní dovednosti. Tato část je zpracována nejpodrobněji. Přesto je možné s některými požadavky polemizovat, například:

- „V poslechu s porozuměním... žák rozliší citové zbarvení, styl...; ...vhodně reagovat na partnerovy podněty“ (Katalog požadavků, 2000, s. 6). Rozlišit citové zbarvení patří k dovednostem osvojeným na vyšších jazykových úrovních; proto je tento požadavek pro úspěšné složení maturitní zkoušky příliš vysoký. Co se týká požadavku „vhodně reagovat na partnerovy podněty“, lze tuto dovednost pokládat spíše za dovednost produktivní, resp. interaktivní, než za dovednost receptivní.
- „Ve čtení s porozuměním ...žák umí rozlišit ...styl, ...postihnout logickou strukturu textu“ (Katalog požadavků, 2000, s. 6). Podobně jako ve výše zmíněném bodu jde o požadavek, který odpovídá spíše vyšší jazykové úrovni.

### 7.3.3 Produktivní dovednosti

Přestože se Katalog zaměřuje na požadavky potřebné ke složení společné části maturity, domnívám se, že v dokumentu tohoto druhu by neměly chybět ani požadavky na školní část maturity. Určitě by bylo vhodné, aby byl pro hodnocení produktivních řečových dovedností formulován alespoň určitý společný rámec, který by se týkal typů testových úloh, jakož i kritérií hodnocení. Pokud bude hodnocení ústního projevu plně v kompetenci jednotlivých středních škol, jak Katalog stanovuje (Katalog požadavků, 2000, s. 5), a tento rámec formulován není, hrozí riziko, že ústní část maturitní zkoušky bude vedena stejným způsobem, jako tomu bylo dosud. Kromě toho lze očekávat, že zkouška ovlivní výuku tak, že důraz bude kladen na rozvoj receptivních řečových dovedností, které jsou předmětem společné části maturity a tudíž získané výsledky na různých školách budou moci být srovnávány.

Formulace týkající se požadované úrovně produktivních řečových dovedností a uvedení doporučených testovacích technik jsou v Katalogu nedostatečné; na druhé straně seznam požadavků na zvládnutí jazykových prostředků je až příliš podrobný – věnuje se mu více než pět stránek, zatímco části popisující požadované řečové dovednosti pouze stránka jedna. Tato skutečnost určitě nebude signálem pro učitele ani žáky, že znalost jazykových prostředků nemůže být při učení a vyučování cizího jazyka cílem, ale pouze prostředkem ke splnění cíle (komunikace v cizím jazyce).

Ve světle těchto skutečností je nutno mít na paměti, že u současné ústní maturitní zkoušky často dochází k tomu, že se žáci „naučí“ (namemorují) určitá vypracovaná „témata“, která pak s větším nebo menším úspěchem před komisí zreprodukují. Při nedostatečném důrazu na ústní zkoušku a nedostatečné formulaci požadavků, testovacích metod a způsobu hodnocení je možné, že tento přístup k ústnímu testování bude nadále přetrvávat. Tento způsob zkoušení však nemá nic společného se zjišťováním úrovně dovednosti mluvení – k tomu je zapotřebí používat zcela odlišné nástroje. Jinými slovy, současný způsob zjišťování úrovně mluvení postrádá jak jasný a srozumitelný popis dovedností (co žák umí říct nebo vyjádřit), tak doporučené testovací techniky a v neposlední řadě také kritéria hodnocení (co a jak dobře žák na určité úrovni umí ústně vyjádřit). Pokud by učitelé středních škol vedli školní část maturity sami, bylo by nutné poskytnout jim nejenom jasné a srozumitelné pokyny k tomu, jak vést ústní zkoušku a jak ji hodnotit, ale také praktické proškolení. V opačném případě by nebylo garantováno

používání srovnatelných testovacích metod a kritérií hodnocení na různých školách. Pouhé vydání a distribuce určitých pokynů nezajišťuje automaticky stejný přístup k administraci i hodnocení ústního projevu. Kromě jiných opatření je nezbytné dát učitelům i možnost vidět ukázky vedení zkoušky a hodnocení kandidáta, v ideálním případě poskytnout pro všechny zkoušející odpovídající školení.

Protože Katalog neuvádí doporučené techniky ústního testování, můžeme pouze odhadovat problémy, které by při zkoušení i hodnocení ústního projevu mohly vzniknout. Instituce odpovídající za standardizované zkoušky z anglického jazyka, které se administrují na celém světě, se neomezují na uvádění tematických okruhů. Zaměřují se na to, aby kandidáty jasně informovaly o formátu testu (subtesty pro jednotlivé dovednosti, délka, typy testových úloh). Kandidát pro svou představu o průběhu ústní zkoušky dostane všechny potřebné informace. Například u zkoušky KET (Key English Test) se na internetových stránkách kandidát dočte, že v části zkoušky mluvení se bude hodnotit jeho/její schopnost interakce, schopnost ptát se a odpovědět na otázky a mluvit o sobě (<http://www.cambridgeesol.org/exams/ket.htm>). Podobně u zkoušky FCE (First Certificate in English) se u popisu formátu testu žádné tematické oblasti neuvádějí, ale kandidát se může seznámit s tím, jak bude test probíhat a co se u ústního projevu hodnotí. Zkouška podle informací na internetu hodnotí opět schopnost interakce, a to v různých kontextech. Obsahuje čtyři části - kratší monolog, skupinový úkol a diskusi (<http://www.cambridgeesol.org/exams/fce.htm>).

Návrh zjišťovat úroveň dovednosti psaní jako další produktivní řečové dovednosti je zcela oprávněný. Na druhé straně, pokud nebudou jasně specifikovány požadavky na dosažení určité úrovně, nebudou doporučeny vhodné metody testování ani kritéria hodnocení, opět je možné, že se žáci předem připraví na psaní určitých témat, která si nechají zkontrolovat a poté u písemné zkoušky toto naučené téma napíší. Takovýto způsob testování nám opět neřekne nic o skutečné úrovni psaní u jednotlivých žáků. Podobně jako pro zjišťování úrovně mluvení je důležité určit co nejpřesněji a nejsrozumitelněji požadavky pro určitou úroveň, dále uvést odpovídající typy testových úloh a nakonec také stanovit kritéria hodnocení. Přestože tato doporučení mohou vypadat jednoduše a srozumitelně, je velmi složité je zpracovat tak, aby je žáci, učitelé i hodnotitelé chápali stejným způsobem. Pokud však, podobně jako u dovednosti mluvení, mají učitelé správně chápat požadavky k úspěšnému složení zkoušky z dovednosti psaní, nestačí pouze vyjmenovat požadované gramatické struktury, zviádnuté části syntaxe a rozsah slovní zásoby. Je také potřeba seznámit je s ukázkami úkolů na různých úrovních, dle možnosti také s komentářem, proč byla každá práce hodnocena tímto způsobem. Pokud učitelé sami nehodnotí práce kandidátů, nepotřebují složitý systém školení v této oblasti. Na druhou stranu musejí správně pochopit nejenom požadavky na jazykovou úroveň, ale i principy hodnocení, aby mohli správně připravovat žáky ke zkoušce. Jinými slovy, je zcela nedostačující pouze vyjmenovat oblasti gramatiky, syntaxe a slovní zásoby; pokud není stanoveno, jak dobře, v jakých situacích a za jakým účelem je potřeba uvedené znalosti aplikovat, jsou požadavky formulovány nedostatečně.

Tato doporučení lze podpořit také zmínkou o důležitosti hodnotících kritérií u Bachmana a Palmera. Jako příklad uvádějí hodnocení eseje nebo nahrávky ústní zkoušky, kdy



hodnotitel přibližně ví, na co by se měl při hodnocení koncentrovat, avšak „pokud hodnotitelům není přesně řečeno, na které oblasti jazykových schopností se mají zaměřit, budou se jejich hodnotící kritéria lišit“ (Bachman; Palmer 1996, s. 202-203).

Katalog dále uvádí některé požadavky na produktivní řečové dovednosti, se kterými nelze jednoznačně souhlasit.

- „V písemném projevu žák zná a umí srozumitelně, v souladu s pravopisnými normami a stylisticky vhodně ...napsat...poznámky k přečtenému nebo vyslechnutému textu (osnovu, anotaci, teze, resumé), ...napsat zápis z jednání či besedy,...“. Psaní poznámek k vyslechnutému textu totiž neumožňuje pouze dovednost psaní, ale také dovednost poslechu. Psaní poznámek se považuje za nezbytnou dovednost například pro studenty studující v zahraničí. Za tímto účelem se publikují i speciální učebnice s cílem připravit zahraniční studenty na studium na univerzitách v anglicky mluvících zemích. Jako příklad je možné uvést učebnici *Study listening*, jejíž součástí jsou lekce 6 a 12, které se zaměřují právě na to, jak se naučit dělat si poznámky v cizím jazyce (Lynch, 1983, s. 19, s. 51). Dovednost psaní poznámek z mluveného textu samostatně uvádí také Společný evropský referenční rámec, dále SERR (Common European Framework, 2001, s. 96). Psaní poznámek v cizím jazyce je nepochybně dovednost užitečná a praktická. Otázkou je, jestli je možné věnovat přípravě na ni ve středoškolském studiu dostatečný čas a pozornost a také jestli jsou učitelé jazyků schopni tuto přípravu vést. Kromě toho by bylo potřebné jasně specifikovat kritéria hodnocení, protože je často velmi komplikované odlišit, jestli je příčinou nízké úrovně psaní poznámek problém kandidáta v poslechu s porozuměním nebo v psaní samotném.
- „...V písemném projevu žák umí ...psát podle diktátu“ (Katalog požadavků, 2000, s. 6); „v ústním projevu žák umí jazykově správně, srozumitelně a přiměřeně situaci ... přednést nazpaměť naučený text (báseň, píseň apod.), číst nahlas“, (Katalog požadavků, 2000, s. 6). Za diskutabilní je možné označit požadavky „psát podle diktátu“ nebo „číst nahlas“. Obě testovací techniky se používaly v jazykovém testování spíše v minulosti, nicméně některá odborná literatura se o nich stále zmiňuje. Například *Dictionary of Language Testing* uvádí, že testovací technika diktát hodnotí více dovedností, které však dále nespecifikuje. Podle tohoto zdroje použití techniky diktátu umožňuje dobré rozlišení kandidátů podle schopností. Nepokládá však tuto techniku testování za dostatečně validní. Za problematické a obtížné pokládá navíc zkonstruování vhodných hodnotících škál (Davies et al., 1999, s. 44). Jiní autoři zase poukazují na skutečnost, že tuto metodu je možné spravedlivě použít pouze tehdy, když je diktát prezentován všem kandidátům stejným způsobem. Jako další nevýhodu uvádějí obtížnost nejenom zkonstruování hodnotících škál, ale také obtížnost a časovou náročnost opravy diktátů, jakož i vysokou pravděpodobnost, že při opravě dojde k mnohým chybám (Alderson; Clapham; Wall 1995, s. 57). Techniku čtení nahlas uvádí Underhill a kromě některých výhod uvádí také nevýhody, jako například, že tato technika není autentická ani komunikativní. Navíc ústní projev kandidáta při použití této techniky závisí do velké míry na dovednosti čtení (Underhill, 1994, s. 76).

#### 7.3.4 Mediační dovednosti

Mediační dovednosti se u běžných mezinárodně uznávaných jazykových zkoušek nepožadují, a to z pochopitelných a velmi praktických důvodů – tyto zkoušky jsou určeny kandidátům z celého světa a tedy zpracovávat dvojjazyčné testové úlohy pro jednotlivé kandidáty by nebylo možné. Neznamená to sice automaticky, že by ověřování mediačních dovedností nemohlo nebo nemělo být součástí maturitní zkoušky z jazyka. Avšak často jde o dovednosti, které je dost složité osvojit si na střední škole a v neposlední řadě je také těžké je spolehlivě hodnotit.

Katalog formuluje požadavek, aby žák uměl přeložit text. Překlad textu je činnost velmi komplexní a nevypovídá pouze o úrovni znalostí a dovedností v cizím jazyce, ale také o schopnosti adekvátně přetransformovat text z jednoho jazyka do druhého. Katalog neuvádí typ textů, které by student měl umět přeložit, jejich rozsah, úroveň obtížnosti, ani jestli se má jednat o překlad s použitím nebo bez použití slovníku. Navíc se jedná o specifickou dovednost, kterou při běžném učení se jazyku není možné dostatečně si osvojit. Velmi složitou by pak byla také otázka hodnocení překladu, zejména zajištění konzistentnosti hodnocení různými hodnotiteli. Další možnou komplikací by byla odpověď na otázku, zda se v případě překladu z cizího jazyka do jazyka mateřského hodnotí porozumění cizojazyčnému textu, nebo také způsob, jak je obsah přetransformován do mateřského jazyka; u překladu navíc neexistuje pouze jeden jediný možný správný překlad. Odborná literatura zabývající se testovacími technikami většinou překlad nebo tlumočení neuvádí, případně je zmiňuje pouze jako techniky okrajové. Fulcher uvádí typy testových úloh, kde kandidát sehrává roli tlumočnicka (Fulcher, 2003, s. 83 - 84). Underhill, přestože tuto testovací techniku uvádí, se domnívá, že se jedná o techniku zastaralou a nevhodnou pro výukové jazykové programy používající přímou metodu. Navíc zastává názor, že dovednost dobře překládat je odlišná od dovednosti dobře mluvit cizím jazykem. Tuto testovací techniku nepokládá za dostatečný test mluvních dovedností (Underhill, 1994, s. 79).

#### 7.3.5 Tematické okruhy

Kromě požadavků na jednotlivé řečové dovednosti Katalog uvádí i tematické okruhy, v jejichž rámci má žák u maturitní zkoušky prokázat úroveň všeobecných i komunikativních kompetencí. Několikrát je uvedeno, že tematické okruhy nejsou samy o sobě předmětem ověřování společné části maturity. Na druhé straně některé z tematických okruhů jakoby přímo nabádaly k tomu, jak se na ústní část maturity „připravit“ nebo „naučit“, například: Můj domov (popis), Naše škola (popis, „připravit“ nebo „naučit“, například: Můj domov (popis), Naše škola (popis, charakteristika předmětů, rozvrh hodin), Zaměstnání (volba povolání, druhy povolání, mzda, nezaměstnanost) a jiné. Jiná témata zase předpokládají i věcnou znalost problematiky, přestože tematické okruhy nemají být podle katalogu předmětem ověřování. Jedná se například o tematický okruh Člověk a společnost, který obsahuje kromě jiného tyto body: Vztahy mezi lidmi (nářadí, láska, rasismus; vztah k postiženým, cizincům, menšinám); Postoje, názory (např. situace ve společnosti – k politice, ekonomice, etice, životní styl) atd. (Katalog požadavků, 2000, s. 7). Na tato témata není možné mluvit bez určité znalosti problematiky ani v mateřském jazyce, natož

v jazyce cizím. Pokud by kandidát neuměl o těchto tématech komunikovat, bylo by otázkou, jestli jde o problém jeho neznalosti faktických informací nebo o neschopnost ústní komunikace. Další možnou spornou otázkou by pak bylo, jestli by zkoušející hodnotili v ústním projevu samotný obsah (znalosti) nebo jazykovou úroveň.

Po přečtení seznamu tematických okruhů v Katalogu je vysoce pravděpodobné, že pokud nebudou všechny zúčastněné strany srozumitelně poučeny o cíli zkoušky z dovednosti mluvení, budou za nejlepší přípravu ke zkoušce považovat napsání a naučení se těchto témat. Od zkoušky pak budou pravděpodobně žáci očekávat, že jim bude poskytnuta možnost na jedno z naučených témat přednést svůj předem připravený a z paměti naučený ústní monolog. Komunikativní pojetí zkoušek zjišťujících úroveň dovednosti mluvení je však, jak již bylo zmíněno, poněkud odlišné.

#### 7.4 Jsou formulace v Katalogu požadavků jasné a srozumitelné?

Pokud je Katalog opravdu určen, jak se uvádí v úvodu, žákům, jejich rodičům a učitelům středních škol, měl by být napsán s ohledem na své adresáty a tedy jazykem srozumitelným i laické veřejnosti. Dokumenty informující o zkouškách se v testovacích institucích často vytvářejí v několika odlišných verzích pro různé publikum a podle toho se samozřejmě formálně i obsahově liší. Jinak bude napsán dokument určený tvůrcům testů, ministerským úředníkům, učitelům, kteří se mohou nebo nemusí podílet na tvorbě, zadávání nebo hodnocení testů, nebo dokument určený žákům a rodičům. Obecně se od podobného dokumentu určeného kandidátům, tedy v tomto případě žákům, očekává, že jim poskytne o testu dost informací k tomu, aby mohli podat co nejlepší výkon (Alderson; Clapham; Wall 1995, s. 21).

Samotná struktura Katalogu je nepřehledná – první část CÍLOVÉ KOMPETENCE uvádí cílové kompetence, které se mají realizovat „prostřednictvím řečových dovedností a na základě jazykových prostředků“. Domnívám se, že lepší formulací by bylo „prostřednictvím řečových dovedností a na základě znalostí jazykových prostředků, které zahrnují znalost gramatiky, syntaxe jazyka a slovní zásoby“. V části CÍLOVÉ KOMPETENCE jsou uvedeny podkapitoly Řečové dovednosti a Jazykové prostředky. Poté se uvádějí KOMUNIKAČNÍ SITUACE A TEMATICKÉ OKRUHY. Další částí Katalogu jsou SPECIFICKÉ CÍLE, které víceméně opakují obsah podkapitoly Řečové dovednosti a rozšiřují je pouze o typy a délku textů, které mohou být v testech použity. Také rozlišují popisy požadavků pro základní a vyšší úroveň. Avšak popis rozdílů v požadavcích je dost vágní; například jeden z požadavků na základní úroveň (Z) je, že „...žák poslech s porozuměním a čtení s porozuměním pro základní úroveň (Z) je, že „...žák rozumí ...sdělení ...s velmi malým počtem neznámých slov“, zatímco pro úroveň vyšší (V) je požadavek, že „...žák rozumí ...sdělení ...s malým počtem neznámých slov“. U specifických cílů pro čtení s porozuměním se předpokládá, že „žák rozezná záměr i postoj autora, rozezná funkční styl textu“. Ve stejné podkapitole se také uvádí, že žák „umí využívat ilustrací, tabulek“, pro vyšší úroveň také „schémat“. Domnívám se, že autoři měli spíše na mysli to, že žák ilustracím, tabulkám a schématům v cizojazyčném textu rozumí, umí v nich najít požadované informace nebo údaje, případně tyto informace a údaje dále zpracovat. Část SPECIFICKÉ CÍLE, jak už bylo uvedeno, pojednává

o receptivních dovednostech, avšak k produktivním dovednostem se uvádí pouze jedna věta vztahující se k písemnému projevu: „žák zná a umí srozumitelně, v souladu s pravopisnými normami a stylisticky vhodně používat vybrané jazykové prostředky.“ Tento popis požadavků, respektive podle Katalogu popis specifických cílů, je zcela nedostatečný. Požadavky na úroveň dovednosti psaní se běžně specifikují daleko podrobněji. Navíc požadovat víceméně bezchybné používání jazykových prostředků i pro základní úroveň nepokládám za správné. Protože specifické cíle vůbec neuvádějí, co má žák umět v dovednosti mluvení, pokládám tuto část za nedostatečně zpracovanou a tudíž ne srozumitelnou. Ani žáci, ani učitelé a ani rodiče na základě tohoto dokumentu nebudou mít jasnou představu, co má žák v těchto podle mého názoru velmi důležitých řečových dovednostech prokázat, aby byl u zkoušky úspěšný.

Přestože u řečových dovedností v kapitole o cílových kompetencích se uvádějí i mediační dovednosti, kapitola Specifické cíle už tyto dovednosti vůbec neuvádí. Pokud se tedy autoři rozhodli, že tyto dovednosti do požadavků patří, bylo by vhodné, aby byly dále rozvedeny v kapitole Specifické cíle.

## 7.5 Další otázky související s Katalogem požadavků

**Které dovednosti by se měly zkoušet u maturitní zkoušky? Je správné testovat znalost gramatiky odděleně?**

Při úvahách o tom, co by měly testovat maturitní zkoušky, je potřeba začít se stanovením účelu testu. Účelem testu může být například prověřit, jak dobře si žáci osvojili středoškolské učivo z anglického jazyka nebo zjistit, jestli žáci dosáhli určité požadované úrovně, která je stanovena centrálně. Je možné ztotožnit se s názorem Hughese, který sice souhlasí s obecným rozdělením testů na testy dosažené úrovně (*proficiency*) a na testy ověřující zvládnutí učiva po absolvování uceleného výukového kurzu (*end of course achievement test*), ale zároveň nesouhlasí s použitím odlišných procesů při určování obsahu kurzů. Svůj názor podporuje tím, že testy zadávané na konci kurzu by měly spíše než z obsahu kurzu vycházet z cílů kurzu, čímž budou velmi podobné, ne-li stejné, jako testy dosažené úrovně (Hughes, 1988 in Alderson; Clapham; Wall 1995, s. 23).

Protože maturitní zkouška není zkouškou odbornou a předpokládá se, že bude zadávána na všech typech středních škol, měla by být spíše obecného charakteru a ověřovat řečové dovednosti, které jsou potřebné v běžném životě a které zkoušky uznávané mezinárodně také testují. Nabízí se tedy možnost testovat čtyři základní řečové dovednosti (poslech s porozuměním, mluvení, čtení s porozuměním, psaní), případně také separátní část testující znalost jazykových prostředků (gramatika, syntax, stylistika, slovní zásoba). Zařazení takovéto části má své výhody i nevýhody. Pokud se do testu část zjišťující úroveň znalostí jazyka zařadí, bude se tato část testu konstruovat jednodušeji a také hodnocení bude poměrně snadné ve srovnání s hodnocením řečových dovedností. Zavedení této části však může také zpětně ovlivnit výuku tak, že se bude klást důraz na zavedení této části však může také zpětně ovlivnit výuku tak, že se bude klást důraz na vyučování gramatiky a slovní zásoby spíše než na rozvoj řečových dovedností. Ve výuce cizích jazyků v České republice byla totiž tradiční gramaticko-prekladová metoda aplikována jako metoda hlavní a tyto přístupy na všech stupních školství často ještě stále přetrvávají. Proto v případě, že by maturitní zkouška část týkající se užívání jazyka

neobsahovala, je možné, že by se ze strany učitelů objevily negativní reakce; zároveň je však možné, že by se rozvoji základních řečových dovedností při vyučování cizím jazykům začala věnovat potřebná pozornost. Když se pro srovnání podíváme na standardizované mezinárodně uznávané jazykové zkoušky, zjistíme, že zatímco jazykové zkoušky pro nižší úroveň (PET, KET) obsahují pouze dílčí testy ověřující základní řečové dovednosti, jazykové zkoušky vytvořené pro ověření vyšších úrovní (FCE, CAE ) zahrnují také část zaměřenou na zjišťování úrovně používání jazyka (tedy zejména správného a přesného používání gramatických pravidel a slovní zásoby), viz. ČÁST I, tab. 2. Protože znalost jazykových prostředků a jejich správné používání by mělo být na vyšší úrovni řečových dovedností samozřejmostí, je její ověření dílčím testem vhodné a obhajitelné.

### **Jaký dopad na výuku cizích jazyků může mít zavedení státní maturitní zkoušky?**

Protože maturitní zkoušce se v naší zemi obecně přikládá velká důležitost, je možné předpokládat, že k nové maturitní zkoušce budou všechny zainteresované strany přistupovat s vážností a velmi zodpovědně a že maturitní zkouška pravděpodobně zpětně ovlivní výuku (viz ČÁST I, subkapitola 3.2). Přestože tento efekt není přímočarý a uvádí se pouze málo empirických důkazů o jeho existenci, při návrhu maturitní zkoušky by nebylo vhodné možnost jeho vzniku opomínat. Test by proto měl mít takové vlastnosti, o kterých by se dalo předpokládat, že vyvolají pozitivní zpětný efekt. V případě, že by se testovaly pouze receptivní dovednosti, kterými se Katalog zabývá, a zjišťování úrovně dovedností produktivních by bylo zcela v kompetenci školy, s největší pravděpodobností by se výuka zaměřila na procvičování dovedností poslechu a čtení s porozuměním, protože právě tyto dovednosti by byly testovány ve společné části maturity a je možné, že by se výsledky žáků různých škol srovnávaly. Proto by se snahy žáků i učitelů mohly zaměřit právě na tyto dovednosti, což by rozhodně nebylo správné. Rovněž rozsáhlý seznam požadovaných znalostí jazykových prostředků by mohl způsobit, že by se výuka na tuto oblast nepřiměřeně koncentrovala.

Z uvedených důvodů jsem přesvědčena, že zpracování Katalogu je činnost velmi zodpovědná a náročná a že Katalog by měl jasně formulovat požadavky na úroveň, metody testování a kritéria hodnocení pro všechny řečové dovednosti.

## **8. Návrh alternativního Katalogu požadavků pro státní maturitní zkoušku z anglického jazyka**

### **8.1 Katalog požadavků pro žáky, rodiče a širokou veřejnost**

Na základě poznatků získaných jak při studiu odborné literatury a při psaní první části této práce, tak i na základě seznámení se s problematikou maturitní zkoušky z cizího jazyka obecně a s Katalogem požadavků konkrétně byl vytvořen nový, alternativní Katalog požadavků pro maturitní zkoušku z anglického jazyka.

#### **Účel maturitní zkoušky z cizího jazyka**

Účelem maturitní zkoušky z anglického jazyka je ověřit jazykové znalosti a řečové dovednosti kandidátů a poskytnout platné a spolehlivé osvědčení o jejich úrovni. Nejedná se o zkoušku, která zjišťuje zvládnutí určité konkrétní jazykové učebnice, gramatických pravidel nebo naučení se jistého množství tématických okruhů či faktografických údajů týkajících se kultury a/nebo historie anglicky mluvících zemí. Zkouška nemá za cíl hodnotit paměť ani inteligenci kandidátů. Zkouška hodnotí komplexně úroveň základních řečových dovedností poslechu, čtení, psaní a mluvení, pro úroveň středně pokročilou hodnotí zkouška také užití jazyka. U studentů, kteří dosahují této úrovně, se totiž očekává přesnější používání jazyka, a proto je vhodné zařadit samostatnou součást testu, která ho ověřuje. Dosažené výsledky v jednotlivých řečových dovednostech budou vyjádřeny jazykovým profilem kandidáta, který pro mírně pokročilou úroveň obsahuje 4 známky (poslech, čtení, psaní a mluvení) a pro středně pokročilou úroveň 5 známek (poslech, čtení, psaní, užití jazyka a mluvení). Například žák, který absolvoval zkoušku pro mírně pokročilou úroveň a dosáhl v testu z poslechu s porozuměním známky 2, ze čtení s porozuměním 2, ze psaní 1 a z mluvení 3, bude mít svůj jazykový profil 2213. Takto vyjádřeny výsledky zkoušky podávají totiž přesnější obraz o dosažené úrovni jednotlivých řečových dovedností žáků, než pouze jedna známka. V případě potřeby vyjádřit hodnocení jednou známkou je možné vypočítat aritmetický průměr. V uvedeném případě by byl žák hodnocen celkově známkou 2. Tato jediná známka však neposkytuje přesnou informaci například o tom, že dovednost mluvení je horší než známka dvě, zatímco dovednost psaní je naopak lepší než známka dvě.

#### **8.1.1 Požadovaná úroveň řečových dovedností**

Maturitní zkouška nabídne kandidátům možnost výběru ze dvou rozdílných úrovní obtížnosti, a to mírně a středně pokročilé. Mírně pokročilá úroveň přitom vychází z požadavků na úroveň A2 podle Společného evropského referenčního rámce (SEER), z požadavků na úroveň A2 podle Společného evropského referenčního rámce (SEER), středně pokročilá pak vychází z požadavků stanovených v popisech úrovně B2 podle SEER.

Společný evropský referenční rámec pro jazyky, jak už byl zmíněno v úvodu práce, je evropský dokument detailně popisující 6 jazykových úrovní od nejnižší A1 po nejvyšší C2. Jeho účelem je zjednodušit vzájemné porozumění mezi různými evropskými

vzdělávacími institucemi zabývajícími se výukou a hodnocením jazyků. Tento dokument je běžně dostupný, např. na internetu v anglické verzi

[http://www.coe.int/T/E/Cultural\\_Cooperation/education/Languages/Language\\_Policy/Common\\_Framework\\_of\\_Reference/1cadre.asp#TopOfPage](http://www.coe.int/T/E/Cultural_Cooperation/education/Languages/Language_Policy/Common_Framework_of_Reference/1cadre.asp#TopOfPage)

<http://culture2.coe.int/portfolio/documents/0521803136txt.pdf>

nebo v české verzi

<http://www.msmt.cz/DOMEK/default.asp?ARI=101945&CAI=2968>.

Podle SERR je kandidát na úrovni A2 (uživatel základů jazyka) charakterizován následujícím způsobem: „Rozumí větám a často používaným výrazům vztahujícím se k oblastem, které se ho/jí bezprostředně týkají (např. základní informace o něm/ní a jeho/její rodině, o nakupování, místopisu a zaměstnání). Dokáže komunikovat prostřednictvím jednoduchých a běžných úloh, jež vyžadují jednoduchou a přímou výměnu informací o známých a běžných skutečnostech. Umí jednoduchým způsobem popsat svou vlastní rodinu, bezprostřední okolí a záležitosti týkající se jeho/její nejnáléhavějších potřeb.“

([http://culture2.coe.int/portfolio/documents/globalscale\\_czech.doc](http://culture2.coe.int/portfolio/documents/globalscale_czech.doc))

Kandidát na úrovni B2 je podle SERR charakterizován následovně: „Dokáže porozumět hlavním myšlenkám složitých textů týkajících se jak konkrétních, tak abstraktních témat včetně odborně zaměřených diskusí ve svém oboru. Dokáže se účastnit rozhovoru natolik plynule a spontánně, že může vést běžný rozhovor s rodilými mluvčími, aniž by to představovalo zvýšené úsilí pro kteréhokoliv účastníka interakce. Umí napsat srozumitelné podrobné texty na širokou škálu témat a vysvětlit své názorové stanovisko týkající se aktuálního problému s uvedením výhod a nevýhod.“

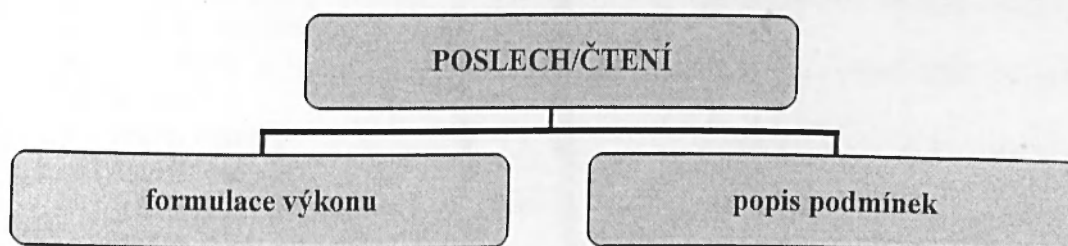
([http://culture2.coe.int/portfolio/documents/globalscale\\_czech.doc](http://culture2.coe.int/portfolio/documents/globalscale_czech.doc))

### 8.1.2 Doporučení vhodných učebních materiálů

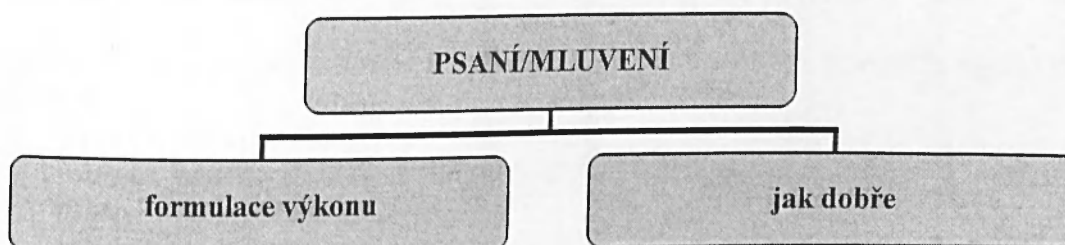
Pro přípravu k maturitní zkoušce není doporučen jeden konkrétní výukový materiál. Záleží na preferencích jednotlivých škol či učitelů. Obecně platí, že výukový materiál by měl být jedním z britských nebo amerických jazykových kurzů v současné době běžně komerčně dostupných na knižním trhu. Dle výstupních cílů konkrétních tříd nebo jazykových skupin je pak vhodné zvolit buď kurz pro mírně nebo středně pokročilé (*pre-intermediate* nebo *intermediate*). Většina těchto učebnic nabízí vyvážený materiál obsahující cvičení zaměřená na vyrovnaný rozvoj všech řečových dovedností. Zkušený učitel však používání učebnice přizpůsobuje „na míru“ konkrétním skupinám žáků a doplňuje je vhodně i jinými, doplňkovými materiály. Doporučuje se používat také různé autentické materiály (např. zdroje dostupné na internetu, brožury, letáky apod.). Při výuce je důležitý nejenom samotný materiál, ale také metody výuky. Je potřeba klást důraz na komunikaci v cizím jazyce; není tedy vhodné používat výukové materiály využívající komunikativní přístupy a současně aplikovat starší gramaticko-překladové metody výuky.

### 8.1.3 Požadavky na výkony v jednotlivých řečových dovednostech

Tento oddíl formuluje požadavky na výkony potřebné k úspěšnému složení zkoušky v jednotlivých dovednostech. Pro receptivní řečové dovednosti (poslech, čtení) na obou úrovních (mírně pokročilá a středně pokročilá) jsou stanoveny požadované výkony a podmínky, za jakých má být výkon plněn (obr. 14). Pro produktivní řečové dovednosti (psaní, mluvení) na obou úrovních jsou stanoveny požadované výkony a také je popsáno, jako dobře musí být popsány výkony provedeny (obr. 15).



Obr. 14. Znárodnění požadavků na receptivní dovednosti



Obr. 15. Znárodnění požadavků na produktivní dovednosti

#### Poslech s porozuměním

*Úroveň mírně pokročilá (A2)*

Požadované dovednosti (výkon)

Kandidát umí při poslechu:

- rozpoznat známé, běžně používané fráze;
- rozpoznat vysokofrekvenční slovní zásobu v jednoduchých větách;
- porozumět hlavnímu sdělení;
- zachytit požadované jednoduché specifické informace (počet, místo, čas, v telefonickém rozhovoru např. jméno a telefonní číslo volaného apod.).



## Podmínky výkonu

Kandidát je schopen výše uvedeného výkonu při splnění těchto podmínek:

- poslech se uskuteční v jazykové laboratoři, s použitím nahrávky a sluchátek a bez použití slovníku;
- nahraný text může být monologický nebo dialogický, mluvený rodilými nebo nerodilými mluvčími; v případě nerodilých mluvčí se nejedná o rodilé mluvčí českého jazyka;
- tempo řeči nahaného textu, ať už produkovaného rodilými nebo nerodilými mluvčími, je modifikováno: je pomalejší než je typické pro autentickou promluvu rodilých mluvčích;
- výslovnost je srozumitelná a poslech není rušen dalšími zvuky nebo okolním hlukem;
- obsah nahrávky se týká každodenních známých a/nebo lehce předvídatelných situací běžného života doma, ve škole a v omezené míře na veřejnosti (jednoduché vzkazy, popisy, instrukce, informace a varování např. na letišti, nádraží;);
- obsah sdělení je konkrétní, faktický (ne abstraktní).

## *Úroveň středně pokročilá (B2)*

### Požadované dovednosti (výkon)

Kandidát umí při poslechu (kromě dovedností požadovaných pro mírně pokročilou úroveň):

- pochopit hlavní myšlenku;
- rozeznat hlavní myšlenku od podpůrných detailů, vyjádření faktu od vyjádření názoru, tvrzení od uvedení příkladu;
- identifikovat hlavní body nahrávky;
- zachytit požadované detailní informace;
- rozpoznat relevantní body textu, které jsou důležité pro shrnutí;
- rozpoznat méně používanou slovní zásobu ve složitějších větách;
- odvodit význam neznámých slov nebo gramatických konstrukcí ze situace (kontextu);
- v omezené míře chápat obsah sdělení „mezi řádky“ (nepřímo vyslovený obsah, tón, postoj mluvčího);
- přenést slyšenou informaci do vizuální nebo grafické formy.

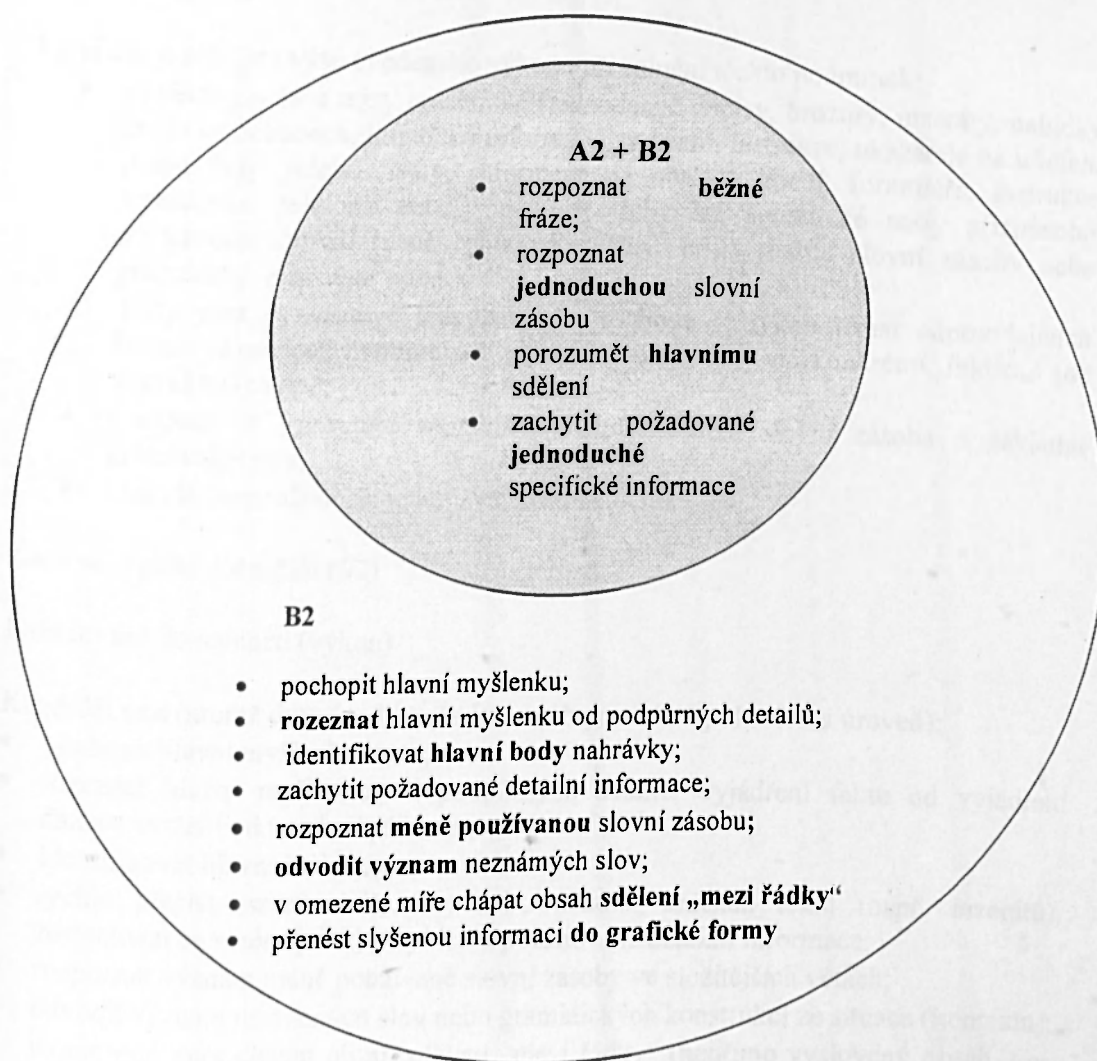
Přestože se předpokládá, že kandidát zvládá i dovednosti požadované pro mírně pokročilou úroveň, jejich ověřování není součástí subtestu úrovně středně pokročilé.

## Podmínky výkonu

Kandidát je schopen výše uvedeného výkonu při splnění těchto podmínek:

- poslech se uskuteční v jazykové laboratoři, s použitím nahrávky a sluchátek a bez použití slovníku;

- nahraný text může být monologický, dialogický nebo obsahující rozhovor více než dvou osob, mluvený rodilým(i) nebo nerodilým(i) mluvčím(i); v případě nerodilých mluvčích se nejedná o rodilé mluvčí českého jazyka;
- tempo řeči se blíží tempu řeči běžnému pro autentický ústní projev rodilých mluvčích;
- výslovnost je srozumitelná, kromě řeči je možné v pozadí slyšet i jiné zvuky nebo okolní hluk;
- obsah nahrávky se týká situací běžného života na veřejnosti, v akademickém prostředí a v omezené míře na možném pracovišti (vybraná vysílání médií – např. zprávy, předpověď počasí, rozhovory, přednášky, prezentace, projevy určené veřejnosti) a je jak konkrétní, tak v omezené míře abstraktní.



Obr. 16. Poslech – požadavky na výkon pro úroveň *mirně pokročilou (A2)* a *středně pokročilou (B2)*

## Čtení s porozuměním

### Úroveň mírně pokročilá (A2)

#### Požadované dovednosti (výkon)

Kandidát při čtení textu v cílovém jazyce umí:

- rozpoznat známé, běžně používané fráze;
- rozpoznat vysokofrekvenční slovní zásobu v jednoduchých větách;
- porozumět hlavnímu sdělení;
- najít požadované jednoduché specifické informace (počet, místo, čas apod.).

#### Podmínky výkonu

Kandidát je schopen výše uvedeného výkonu při splnění těchto podmínek:

- preferovány jsou texty autentické (jednoduché dopisy, brožury, inzeráty, nabídky zboží v obchodech, turistické informace, nařízení, instrukce, ukazatele na ulicích, jízdni řady, jídelní lístky, informace na obalech zboží, formuláře, instrukce k používání telefonu apod.); podle potřeby lze autentické texty přizpůsobit požadované úrovni (např. z hlediska délky, příliš složité slovní zásoby nebo gramatických struktur apod.);
- texty jsou významově jednoznačné, obsahově se týkají témat odpovídajících životní zkušenosti maturanta a pojmově jsou vedeny na konkrétní, faktické (ne abstraktní) rovině;
- v textech se vyskytuje zejména vysokofrekvenční slovní zásoba a základní gramatické jevy;
- kandidát nepoužívá při vykonávání testu slovník.

### Úroveň středně pokročilá (B2)

#### Požadované dovednosti (výkon)

Kandidát umí (kromě dovedností požadovaných pro mírně pokročilou úroveň):

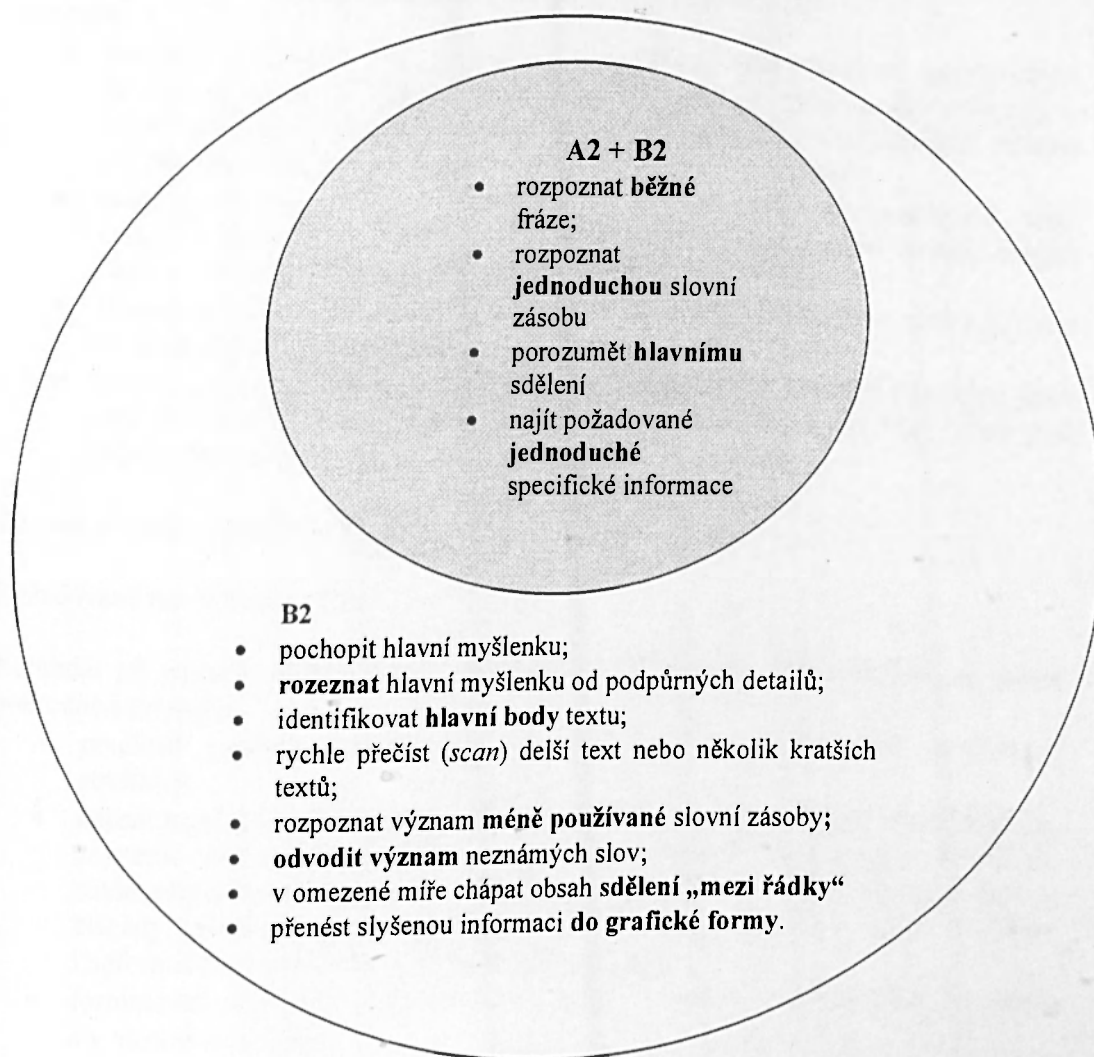
- pochopit hlavní myšlenku textu;
- rozeznat hlavní myšlenku od podpurných detailů, vyjádření faktu od vyjádření názoru, tvrzení od uvedení příkladu;
- identifikovat hlavní body textu;
- rychle přečíst (*scan*) delší text nebo několik kratších textů (např. inzerátů), zorientovat se v něm (nich) a vyhledat požadované detailní informace;
- rozpoznat význam méně používané slovní zásoby ve složitějších větách;
- odvodit význam neznámých slov nebo gramatických konstrukcí ze situace (kontextu);
- v omezené míře chápat obsah sdělení „mezi řádky“ (nepřímo vyslovený obsah, tón, postoj mluvčího).

Přestože se předpokládá, že kandidát zvládá i dovednosti požadované pro mírně pokročilou úroveň, jejich ověřování není součástí subtestu úrovně středně pokročilé.

## Podmínky výkonu

Kandidát je schopen výše uvedeného výkonu při splnění těchto podmínek:

- texty jsou zpravidla autentické; pokud jsou upraveny, jedná se pouze o drobné úpravy;
- texty jsou delší a vyskytují se v nich také méně frekventovaná slova a komplexnější gramatické jevy (novinové články, reportáže, rozhovory, komentáře týkající se současného dění, anotace, grafy, populárně-vědecké články);
- kandidát nepoužívá při vykonávání testu slovník;
- obsah textů se týká situací běžného života na veřejnosti, v prostředí vzdělávací instituce (školy) a v omezené míře na možném pracovišti; obsah textů je jak konkrétní, tak v omezené míře abstraktní.



Obr. 17. Čtení – požadavky na výkon pro úroveň mírně pokročilou (A2) a středně pokročilou (B2)

## Psaní

### *Úroveň mírně pokročilá (A2)*

#### Požadované dovednosti (výkon)

Kandidát při psaní v cílovém jazyce umí:

- formulovat jednoduché myšlenky v krátkých, jednoduchých větách a na konkrétní rovině;
- používat vysokofrekvenční slovní zásobu;
- napsat jednoduchý krátký psaný projev (seznam věcí, vzkaz, přání, dopis apod.);
- vyplnit jednoduché formuláře (žádost, objednávku, informace pro imigrační důstojníky, formuláře v hotelu, při registraci u lékaře apod.)

Jak dobře:

- kandidát prokazuje jistou přesnost v používání jednoduchých gramatických struktur (rozlišení minulého, přítomného a budoucího času), avšak vyskytují se chyby například v tvoření slov, v používání složitějších gramatických struktur a v používání složitějších časů (průběhové a předpřítomné časy);
- může se ale dopouštět chyb i v použití základních jazykových prostředků, např. četných nepřesností v použití i vysokofrekvenční slovní zásoby a také četných chyb v pravopisu;
- schopnost přesných popisů a podání informací je limitována nedostatečnou znalostí jazyka a slovní zásoby;
- tyto limity však nebrání v porozumění podstaty sdělení rodilými mluvčími, kteří jsou zvyklí číst písemný projev cizinců, ani nerodilými mluvčími, kteří jsou zvyklí číst písemný projev cizinců učících se cizímu jazyku;

### *Úroveň středně pokročilá (B2)*

#### Požadované dovednosti (výkon)

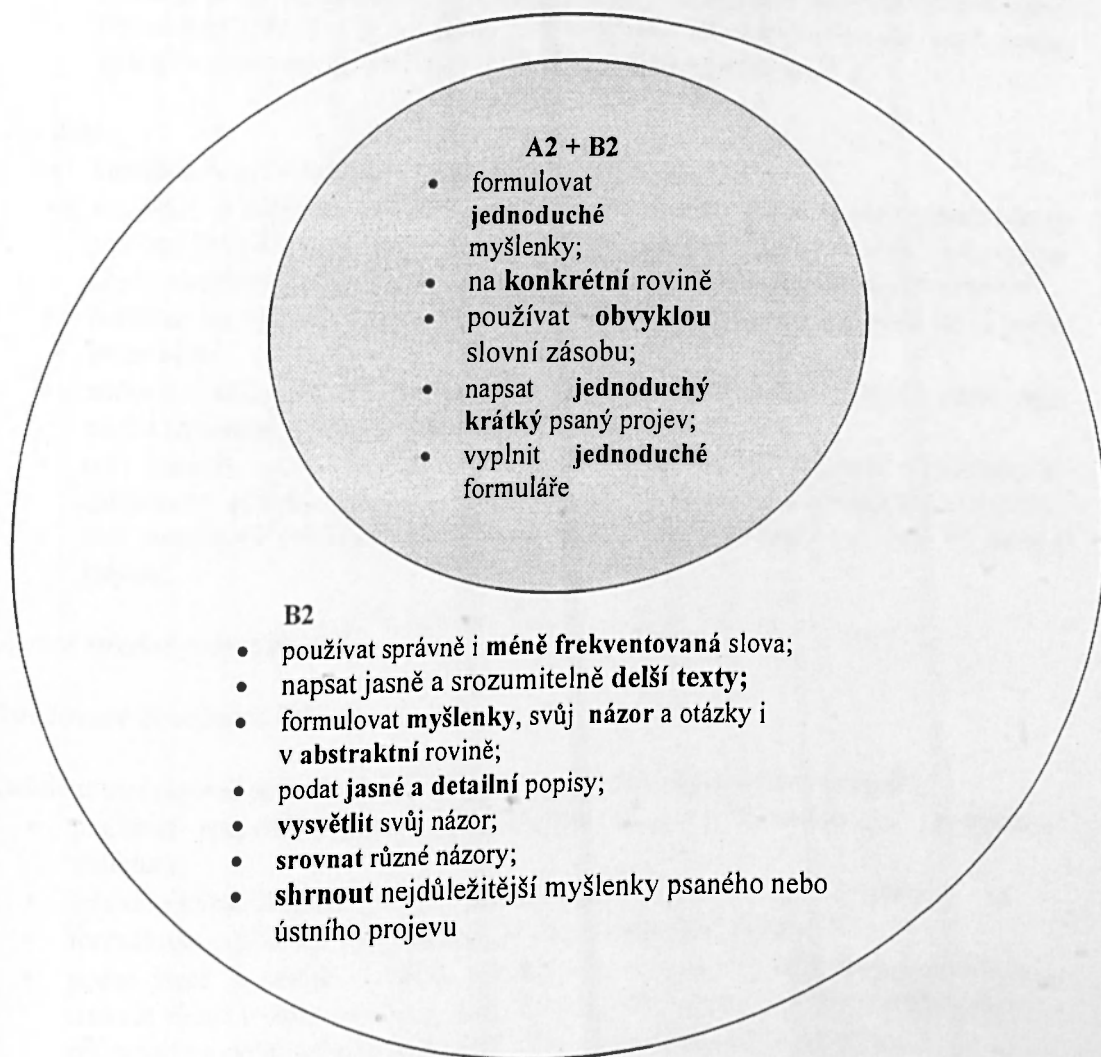
Kandidát při psaní v cílovém jazyce umí (kromě dovedností požadovaných pro mírně pokročilou úroveň):

- používat správně i méně frekventovaná slova a komplexnější gramatické struktury;
- napsat jasně a srozumitelně delší texty pojednávající o široké škále různých témat, zejména však témat kandidátovi známých a blízkých (esej, zpráva, článek do studentských novin, dopis do čtenářské rubriky, žádost o informaci týkající se brigády, studia, práce v zahraničí, strukturovaný životopis, dopis formální i neformální, odpověď na inzerát, stížnost apod.);
- formulovat myšlenky, svůj názor a otázky v delších a složitějších souvětích a v abstraktní rovině;
- podat jasné a detailní popisy, prezentovat a vysvětlit svůj názor/stanovisko, srovnat různé názory na určitý problém a formulovat jejich silné a slabé stránky, zdůraznit ve svém projevu důležité body a uvést podporné argumenty;

Přestože se předpokládá, že kandidát zvládá i dovednosti požadované pro mírně pokročilou úroveň, jejich ověření není součástí subtestu úrovně středně pokročilé.

Jak dobře

- kandidát ovládá a správně používá základní i komplexnější gramatické struktury (např. všechny běžně používané časy, podmínkové věty);
- slovní zásoba se neomezuje pouze na dobře známá vysoce frekventovaná slova;
- kandidát se může dopouštět drobných chyb v používání méně frekventované slovní zásoby a komplexních gramatických struktur;
- chyby v používání komplexnějších gramatických struktur, v přesném používání slovní zásoby i v pravopisu nezpůsobují vážnější nedorozumění v komunikaci s rodilými i nerodilými mluvčími, kteří nejsou zvyklí písemně komunikovat s cizinci a učiteli se anglickému jazyku.



Obr. 18. Psaní – požadavky na výkon pro úroveň mírně pokročilou (A2) a středně pokročilou (B2)

## Mluvení

### *Úroveň mírně pokročilá (A2)*

#### Požadované dovednosti (výkon)

Kandidát při mluvení v cílovém jazyce umí:

- používat běžné a známé fráze;
- používat vysokofrekvenční slovní zásobu v krátkých, jednoduchých větách;
- mluvit o sobě, rodinném životě, škole, každodenních aktivitách, volném čase (jednoduchý popis osoby, prostředí, činnosti; formulování žádosti, instrukcí, otázek; vyjádření toho, co má/nemá rád/a);
- formulovat jednoduché myšlenky v konkrétní rovině;
- získat jednoduché faktické informace potřebné k přežití v anglicky mluvící zemi (formulace jednoduchých otázek, udržování jednoduché konverzace, např. dotazy zjišťující cenu, množství, čas, místo, jak se někam dostat apod. )

#### Jak dobře

- kandidát mluví v krátkých, jednoduchých větách;
- kandidát prokazuje jistou přesnost v používání jednoduchých gramatických struktur (rozlišení minulého, přítomného a budoucího času), avšak vyskytují se chyby například v tvoření slov a v používání složitějších gramatických struktur;
- může se ale dopouštět chyb a nepřesností i v používání jednoduchých jazykových prostředků;
- nedostatečná znalost slovní zásoby může vést k nepřesnostem v popisu osob, věcí, míst a dějů nebo v podání konkrétní informace
- tyto limity spolu s možnou chybou výslovností ale nebrání v porozumění základního sdělení rodilými mluvčími, kteří jsou zvyklí komunikovat s cizinci, ani nerodilými mluvčími, kteří jsou zvyklí komunikovat s učiteli se cizím jazyku;

### *Úroveň středně pokročilá (B2)*

#### Požadované dovednosti (výkon)

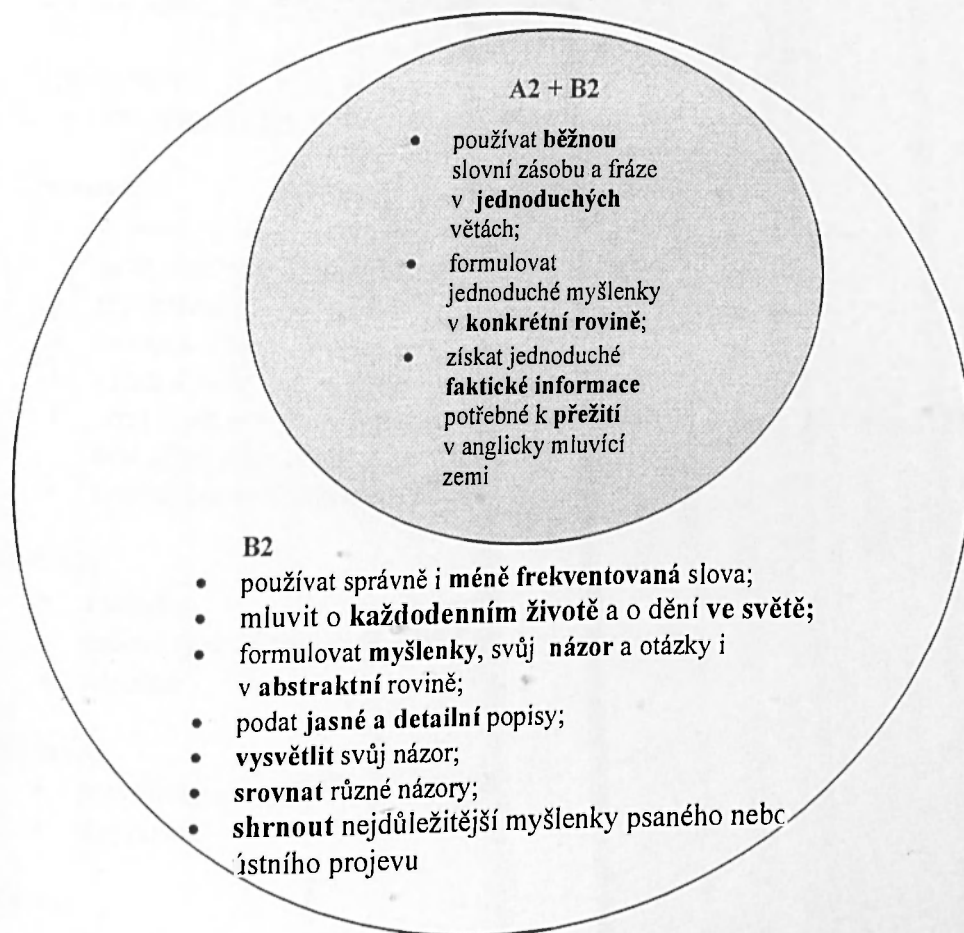
Kandidát umí (kromě dovedností požadovaných pro mírně pokročilou úroveň):

- používat správně i méně frekventovaná slova a komplexnější gramatické struktury;
- mluvit o každodenním životě, o oblastech svého zájmu a o dění ve světě;
- formulovat myšlenky, svůj názor a otázky v abstraktní rovině;
- podat jasné a detailní popisy, prezentovat a vysvětlit svůj názor/stanovisko, srovnat různé názory na určitý problém a formulovat jejich silné a slabé stránky, zdůraznit ve svém projevu důležité body a uvést podpůrné argumenty;
- shrnout ústně nejdůležitější myšlenky psaného nebo ústního projevu.

Přestože se předpokládá, že kandidát zvládá i dovednosti požadované pro mírně pokročilou úroveň, jejich ověřování není součástí subtestu úrovně středně pokročilé.

Jak dobře

- kandidát mluví v delších a složitějších větách;
- kandidát ovládá a správně používá základní i komplexnější gramatické struktury (všechny běžně používané časy, podmínkové věty);
- slovní zásoba se neomezuje pouze na dobře známá vysoko frekventovaná slova;
- kandidát se může dopouštět drobných chyb v používání méně frekventované slovní zásoby a komplexních gramatických jevů; někdy si tyto chyby uvědomí a sám se opraví;
- výslovnost může být nestandardní, avšak je srozumitelná běžným rodilým i nerodilým mluvčím komunikujícím anglicky;
- chyby v používání komplexnějších gramatických struktur, v přesném používání slovní zásoby i ve výslovnosti nezpůsobují vážnější nedorozumění v komunikaci s rodilými i nerodilými mluvčími, kteří nejsou zvyklí na komunikaci s cizinci a učiteli se anglickému jazyku.



Obr. 19. Mluvení – požadavky na výkon pro úroveň mírně pokročilou (A2) a středně pokročilou (B2)



#### 8.1.4 Požadovaná znalost jazykových prostředků

Aby kandidáti mohli prokázat svou úroveň řečových dovedností a splnit požadavky na výkon, musí mít určité znalosti jazykových prostředků, které zahrnují znalosti gramatiky, syntaxe, slovní zásoby a jazykových funkcí.

#### Jazykové znalosti

*Úroveň mírně pokročilá (A2)*

Požadované znalosti o jazyce pro tuto úroveň nejsou samostatně testovány. Jejich seznam se uvádí pouze pro orientaci jak učitelů, tak kandidátů.

#### Podstatná jména

- počítatelná a nepočítatelná (základní rozlišení)
- rod
- jednotné a množné číslo (pravidelné a nepravidelné)
- vyjádření 2. pádu pomocí „of“ a „‘s“

#### Přídavná jména

- stupňování (pravidelné a nepravidelné)

#### Zájmena

- osobní
- ukazovací (*this, that, these, those*)
- přivlastňovací (*my, your...*)
- neurčitá (*some, any* a složeniny)
- vztažná (*who, which, that*)
- tázací (*what, where, when, who, whose, which, why, how, how much, how many, how often, how long*)
- zvrtná (*myself, yourself...*)

#### Číslovky

- základní
- řadové (pravidelné, nepravidelné)
- násobné

#### Příslovce

- pravidelně a nepravidelně tvořené
- stupňování

#### Slovesa

- ve větách oznamovacích, tázacích, v záporu a v imperativu
- základní časy (přítomný a minulý prostý, budoucí s „will“ a „going to“)
- trpný rod (přítomný a minulý čas)

- modální slovesa (základní použití *can, could, would, will, shall, should, may, have, must, mustn't, need, needn't*)
- slovesa *to be, to have*
- základní nepravidelná slovesa
- stavová slovesa

#### Členy

- určitý, neurčitý a nulový (základní použití)

#### Předložky a spojky

- předložky a spojky pro vyjádření základních vztahů mezi slovy ve větě

#### Výrazy

- *there is, there are*
- *much/many, little/few*
- *ago/before*

Tvoření jednoduchých otázek a krátkých odpovědí na ně

Základní pořádek slov ve větě

Nejfrekventovanější frázová slovesa

#### Úroveň středně pokročilá (B2)

Požadované znalosti o jazyce pro tuto úroveň jsou samostatně testovány v části užívání jazyka. Očekává se zvládnutí znalostí o jazyce uvedených pro úroveň mírně pokročilou (A2) a navíc následující znalosti:

#### Slovesa

- všechny běžně používané časy;
- trpný rod ve všech běžně užívaných časech;
- modální slovesa ve všech časech (jejich používání při vyjádření možnosti, určitosti, neurčitosti v minulosti i přítomnosti, např. „*may happen, might have happened, could happen, could have happened, must happen, must have happened...*“)
- nepravidelná slovesa

#### Souvětí s vedlejšími větami:

- podmínkovými (*first, second and third conditional*)
- časovými
- předmětnými
- vztahnými (*defining and non-defining relative clauses*)

#### Používání výrazu

- *used to, be used to, get used to*

### Tvoření otázek

- ve všech časech
- na podmět i přísudek
- zjišťovacích i doplňovacích
- *tag questions*

Nepřímá řeč (*reported speech*)

Nepřímá otázka (*indirect question*)

Vazby sloves s gerundiem a infinitivem

Pořádek slov ve větě

Frázová slovesa

### Slovní zásoba

#### Úroveň mírně pokročilá (A2)

Předpokládá se, že kandidát bude schopen komunikovat ve třech základních oblastech (v závorce jsou uvedeny možné tématické okruhy):

- osobní život (rodina, přátelé, koníčky, volný čas, pocity, názory, zážitky apod.)
- společenský život (nákupy, cestování, prohlídky města, prázdniny, v restauraci, u lékaře, využívání služeb – pošta, banka apod.)
- život ve škole (studium, ubytování, stravování, knihovna, volný čas apod.)

#### Úroveň středně pokročilá (B2)

Kromě slovní zásoby požadované pro úroveň mírně pokročilou (A2) se předpokládá, že kandidát bude schopen komunikovat ve stejných třech základních oblastech na další témata:

- osobní život (rodinné oslavy, návštěvy přátel)
- společenský život (pronájem bytu, auta, vyřízení pojištění, nouzové situace – nehoda, úraz, dopravní přestupek, okradení)
- život ve škole (sledování přednášek a seminářů apod.)

### Jazykové funkce

Tato část podává informaci o tom, co umí kandidát splňující požadavky pro jednotlivé úrovně vyjádřit s použitím uvedených jazykových prostředků a slovní zásoby a za předpokladu, že je schopen požadovaných výkonů v jednotlivých dovednostech. Uvedené funkce se vztahují ve většině případů na dovednosti psaní a mluvení, čili kandidát umí např. „jednoduše popsat lidi, předměty a děje“ písemně i ústně. Některé z funkcí se logicky vztahují pouze na dovednost jednu, např. kandidát umí „požádat o hláskování jména (slova)“ se zjevně vztahuje pouze na dovednost mluvení.

### *Úroveň mírně pokročilá (A2)*

Kandidát umí:

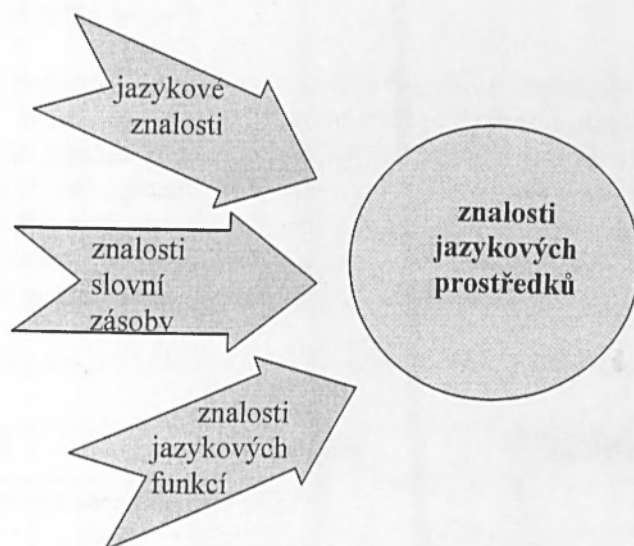
- pozdravit se a odpovědět na pozdrav vhodným způsobem;
- představit se;
- představit své známé a příbuzné;
- vyjádřit souhlas a nesouhlas;
- poskytnout základní osobní údaje (jméno, adresa, datum narození, bydliště, státní příslušnost, vzdělání apod.);
- zeptat se jiné osoby na základní osobní údaje (jméno, adresa, datum narození, bydliště, státní příslušnost, vzdělání apod.);
- jednoduše popsat lidi, předměty a děje;
- požádat o dovolení, zopakování a vysvětlení;
- požádat o vysvětlení neznámého slova;
- požádat o hláskování jména (slova) osobně i po telefonu;
- hláskovat jednoduché osobní údaje o sobě (jméno, příjmení, adresa apod.);
- zeptat se na čas, datum, cenu, jak se dostat na určité místo;
- odpovědět na otázky týkající se vyjádření času, ceny, počtu;
- dát instrukce, jak se dostat na určité místo;
- domluvit si schůzku;
- pozvat na návštěvu, večírek;
- přijmout/odmítnout pozvání;
- vyjádřit, co má rád/a, nerad/a, čemu dává přednost;
- omluvit se a přijmout omluvu.

### *Úroveň středně pokročilá (B2)*

Kandidát umí (kromě toho, co se požaduje pro úroveň mírně pokročilou A2):

- vhodně zahájit, udržovat a ukončit formální i neformální konverzaci;
- popsat přesně lidi, předměty a děje;
- vyjádřit svůj názor a podpořit ho argumenty;
- srovnat dva protichůdné názory, myšlenky;
- požádat o faktickou informaci;
- vyjádřit stížnost;
- poskytnout radu;
- požádat o a nabídnout pomoc;
- vyjádřit, co by se mohlo stát (hypotézy).

Následující obrázek graficky znázorňuje, co obsahují znalosti jazykových prostředků.



Obr. 20. Znárodnění složek znalostí jazykových prostředků

### 8.1.5 Počet dílčích testů a časové limity

Mírně pokročilá úroveň

Pro mírně pokročilou úroveň bude maturitní zkouška z anglického jazyka obsahovat čtyři subtesty, které ověřují poslech, čtení, psaní a mluvení. Časové limity pro tyto dovednosti jsou různé. Testování poslechu je náročné vzhledem k požadované současné koncentraci na poslouchání cizojazyčné nahrávky a odpovědi na otázky, proto je limit stanoven na 20 minut. U dílčího testu čtení, na rozdíl od poslechu, může každý kandidát pracovat svým individuálním tempem, může se k pasážím, které jsou problematické, vrátit a naopak pasáže, které jsou pro něho jednoduché, může vypracovat rychlejším tempem; časový limit je stanoven na 45 minut. V dovednosti psaní na mírně pokročilé úrovni se očekávají spíše jednodušší psané úlohy a tudíž časový limit je kratší, a to 35 minut. Pro ověření úrovně dovednosti mluvení je stanoven časový limit 15 minut.

Tab. 4. Dílčí testy, jejich časové limity a maximální počty bodů pro úroveň mírně pokročilou

Dovednost	Časový limit	Počet bodů
Poslech s porozuměním	20 minut	25
Čtení s porozuměním	45 minut	25
Psaní	35 minut	25
Mluvení	15 minut	25

### Středně pokročilá úroveň

U středně pokročilé úrovně se maturitní zkouška z anglického jazyka bude skládat z pěti subtestů – kromě dovedností, které se ověřují u mírně pokročilé úrovně, se také zjišťuje úroveň užití jazyka. Zatímco kandidáti na mírně pokročilé úrovni se většinou vyjadřují méně přesně, pro splnění požadavků maturitní zkoušky na středně pokročilé úrovni je vhodné ověřit také přesnost užívání jazyka. Ve srovnání s maturitní zkouškou na mírně pokročilé úrovni je maturitní zkouška na středně pokročilé úrovni delší, a to v dílčím testu každé požadované dovednosti, jak ilustruje tabulka.

Tab. 5. Dílčí testy, jejich časové limity a maximální počty bodů pro úroveň středně pokročilou

Dovednost	Časový limit	Počet bodů
Poslech s porozuměním	30 minut	20
Čtení s porozuměním	50 minut	20
Psaní	60 minut	20
Užití jazyka	30 minut	20
Mluvení	20 minut	20

U obou úrovní je maximální počet dosažených bodů 100. U zkoušky na mírně pokročilé úrovni se hodnotí každá ze čtyř základních řečových dovedností stejným počtem bodů 25. U zkoušky na středně pokročilé úrovni se kromě čtyř základních řečových dovedností hodnotí také užití jazyka. Všechny pět subtestů pro tuto úroveň má stejný maximální počet bodů 20.

#### 8.1.6 Testovací techniky

Testovací techniky, které se používají v obou úrovních obtížnosti, jsou více-méně stejné a běžně používané. Testovací techniky se u jednotlivých řečových dovedností liší, resp. testovací techniky používané pro zjišťování úrovně dovednosti čtení a poslechu jsou si často velmi podobné. Testovací techniky používané pro ověřování produktivních dovedností psaní a mluvení se naopak vzájemně liší daleko více. Obecně se dává přednost používání kombinace několika různých testovacích technik. Není nutné, aby každá varianta testu měla totožnou formu. Testovací techniky se tedy mohou verzi od verze lišit. Stejný zůstává stanovený počet bodů pro jednotlivé dovednosti.

Následující typy testovacích technik jsou doporučeny pro ověřování úrovně jednotlivých řečových dovedností; neznámá to však, že v každé verzi jednotlivých subtestů se objeví všechny vyjmenované typy. V každém subtestu jsou použity pouze některé z doporučených technik.

Doporučené typy testovacích technik pro ověřování úrovně *poslechu* (platí pro obě úrovně):

- výběr odpovědi z volby a, b, c nebo d;
- přenos informací z nahrávky do tabulky, formuláře apod;
- pravdivá a nepravdivá tvrzení;
- přiřazování;
- krátké odpovědi na otázky;
- doplňování chybějících informací.

Doporučené typy testovacích technik pro ověřování úrovně *čtení* (platí pro obě úrovně):

- doplňování chybějících informací do textu (grafu, tabulky) dle obsahu textu;
- výběr odpovědi z volby a, b, c nebo d;
- pravdivá a nepravdivá tvrzení;
- přenos informací z textu do tabulky, formuláře apod;
- přiřazování;
- krátké odpovědi na otázky.

Doporučené typy testových úloh pro ověřování úrovně *psaní*

Pro úroveň mírně pokročilou:

- krátký vzkaz, pozdrav, přání apod. dle uvedených instrukcí;
- vyplnění jednoduchého formuláře;
- krátký dopis/email dle uvedených instrukcí;
- popis obrázku;
- odpovědi na otázky.

Pro úroveň středně pokročilou:

- článek do studentských novin;
- dopis do čtenářské rubriky;
- žádost o informaci (týkající se např. brigády, studia, práce v zahraničí) – formou dopisu nebo emailu;
- odpověď na inzerát;
- stížnost.

Doporučené typy testových úloh pro ověřování úrovně *mluvení*

Pro úroveň mírně pokročilou:

- krátké samostatné vyprávění o sobě;
- odpovědi na dotazy zkoušejícího (krátký rozhovor);
- sehrání situační role;
- získání informací a jejich předání.

Pro úroveň středně pokročilou (kromě testových úloh určených pro mírně pokročilou úroveň):

- diskuse se zkoušejícím;
- získání informací od zkoušejícího a jejich přenos jinému zkoušejícímu (vše v anglickém jazyce);

- popis obrázku/série obrázků.

Doporučené typy testovacích technik pro ověřování úrovně *užívání jazyka* (platí pouze pro úroveň středně pokročilou):

- doplňování chybějících slov do textu;
- výběr odpovědi z volby a, b, c nebo d;
- přiřazování;
- seřazování (například vět do správného pořadí).

### 8.1.7 Typy textů

Tato kapitola uvádí typy textů, které se mohou vyskytnout v dílčích testech ověřujících receptivní dovednosti (poslech, čtení).

Typy textů pro poslech

Pro lepší představu kandidátů i učitelů o dílčím testu poslechu uvádíme charakteristiku textů, které mohou být k testování této dovednosti použity pro mírně i středně pokročilou úroveň.

Mírně pokročilá úroveň

Typy poslechových textů jsou gramaticky i lexikálně jednoduché, tempo řeči i výslovnost jsou přizpůsobeny nerodilým mluvčím. Jedná se například o:

- jednoduché vzkazy;
- jednoduchá veřejná oznámení, běžnou veřejnou komunikaci, instrukce (letišť, nádraží, obchody, celnice, pasová kontrola);
- instrukce, jak se kam dostat;
- jednoduché telefonické rozhovory.

Středně pokročilá úroveň

Použité texty jsou obtížnější než texty určeny pro mírně pokročilou úroveň (komplexnější gramatické a syntaktické struktury, méně frekventovaná slovní zásoba, rychlejší tempo řeči, možný hluk v pozadí):

- vzkazy, popisy, instrukce, informace a varování např. na letišti, nádraží;
- veřejná oznámení, běžná veřejná oficiální komunikace, instrukce (letišť, nádraží, obchody, celnice, pasová kontrola);
- instrukce jak se kam dostat;
- telefonické rozhovory;
- rozhlasové, případně televizní zprávy;
- předpověď počasí;
- rozhovory;
- přednášky;
- projevy určené veřejnosti;



- turistické informace.

#### Typy textů pro čtení

Dále uvádíme charakteristiku textů, které mohou být použity pro ověřování dovednosti čtení na obou úrovních pokročilosti.

#### Mírně pokročilá úroveň

Podobně jako u poslechu i u čtení se doporučují texty obsahující vysokofrekvenční slovní zásobu a jednoduché běžné gramatické struktury, např.:

- jednoduché vzkazy;
- jednoduché dopisy;
- informační a turistické brožury;
- inzeráty;
- nabídky zboží v obchodech;
- nařízení pro veřejnost (dopravní, požární);
- ukazatele na ulicích;
- jízdní řády;
- jídelní lístky;
- informace na obalech zboží;
- formuláře;
- instrukce (k používání telefonu, bezpečnostní apod.) ;
- varování.

#### Středně pokročilá úroveň

Použité texty jsou obtížnější než texty určené pro mírně pokročilou úroveň (komplexnější gramatické a syntaktické struktury nebo méně frekventovaná slovní zásoba). Kromě typů textů uvedených pro mírně pokročilou úroveň se mohou vyskytovat v testu následující typy textů:

- recepty;
- programy kin, divadel a jiných kulturní institucí;
- formální dopisy;
- neformální dopisy;
- zpravodajské články v novinách a časopisech;
- populárně-vědecké články;
- manuály k výrobkům denní potřeby (elektrické spotřebiče, výrobky elektroniky) ;
- smlouvy používané v běžném životě (nájemní smlouva, vypůjčení auta).

### **8.1.8 Instrukce k testovým úlohám a kritéria hodnocení**

Instrukce k plnění testových úloh pro všechny dovednosti jsou v anglickém jazyce. Potřebné informace jsou vyjádřeny jasně a srozumitelně.

#### **Kritéria hodnocení písemného projevu**

U písemného projevu se hodnotí:

- splnění cíle komunikace (např.: Je z napsaného jasné, co chtěl pisatel říci? Pochopil by v reálném životě příjemce písemného sdělení, co chtěl pisatel vyjádřit?);
- správnost a přesnost užívání jazyka (dle uvedeného seznamu požadovaných jazykových prostředků pro jednotlivé úrovně);
- obsah sdělení (zejména u delších písemných projevů určených pro středně pokročilou úroveň: Je obsah sdělení logický? Je možné identifikovat hlavní myšlenku a podpůrné argumenty?);
- struktura sdělení (zejména u delších písemných projevů určených pro středně pokročilou úroveň).

Kromě uvedených kritérií byla zpracována podrobnější kritéria, která se vztahují k hodnocení jednotlivých testových úloh ověřovaného subtestu psaní – viz příloha P – 8.

#### **Kritéria hodnocení ústního projevu**

U ústního projevu se hodnotí:

- splnění cíle komunikace (např.: Splnil kandidát úkol? Splnil by kandidát podobný úkol v reálném životě? Podal nebo získal by požadovanou informaci?);
- správnost a přesnost užívání jazyka (dle uvedeného seznamu požadovaných jazykových prostředků pro jednotlivé úrovně);
- úroveň komunikace (např.: Reaguje kandidát na dotazy a/nebo stimuly zkoušejícího? Projevuje v komunikaci iniciativu? Umí se vypořádat se situací, když nerozumí otázce nebo nezná potřebnou slovní zásobu?);
- plynulost ústního projevu;
- výslovnost a intonace.

Kromě zmíněných obecnějších kritérií byla vypracovaná hodnotící škála vztahující se k ukázkovému subtestu mluvení (viz příloha P – 9). Tento subtest ani hodnotící škála nebyly v praxi ověřovány.

## 9. Hodnocení alternativního katalogu požadavků

Poté, co byl vytvořen alternativní katalog požadavků, bylo důležité zjistit, jestli je formulován srozumitelně a jestli se požadavky na jednotlivé úrovně jeví odpovídající. Proto byl zpracován dotazník zjišťující názory učitelů a rodičů na návrh alternativního katalogu. Původně byl určen také k posouzení žáky, avšak z časových důvodů byly nakonec osloveny pouze tři skupiny respondentů – učitelé anglického jazyka na střední škole (4), učitelé anglického jazyka na vysoké škole (3) a rodiče (3). Vzhledem k rozsahu této práce se nejednalo o rozsáhlejší průzkum; šlo pouze o získání alespoň určité zpětné vazby s použitím dotazníku, jakož i neformálních rozhovorů, ve kterých mohli dotázaní vyjádřit své další názory či připomínky.

Dotazník obsahoval položky zaměřující se na obecné i specifické otázky. Otázky obecnějšího charakteru v úvodu dotazníku se týkaly srozumitelnosti formulací, specifické otázky se zaměřovaly na hodnocení náročnosti požadavků pro jednotlivé úrovně, počtů dílčích testů, časových limitů a počtů bodů. Závěr dotazníku obsahoval opět obecnější dotaz týkající se možného dopadu zavedení navrhované maturitní zkoušky na výuku anglického jazyka. Na konci dotazníku byl dotazovaným dán prostor k vyjádření připomínek. Přesné znění dotazníku je uvedeno v příloze P - 3.

Všem dotazovaným byl *úcel i obecná charakteristika* obou požadovaných úrovní srozumitelný. Všichni si také dokáží představit, co se konkrétně od žáků bude u zkoušky požadovat. Všichni dotazovaní se také shodují v tom, že požadavky pro zkoušku na úrovni *mírně pokročilý* mají odpovídající náročnost. V hodnocení požadavků pro zkoušku na úroveň *středně pokročilý* se už názory na náročnost lišily. Zatímco šest z oslovených si myslí, že náročnost je odpovídající (jeden s výhradou, že příliš obtížná je „hypotéza, podpora argumentů a frázová slovesa“), jeden dotazovaný si myslí, že požadavky jsou příliš málo náročné (návrh „doplnit specifikace v písemném projevu, např. linking words a realie“). Dva dotazovaní se naopak domnívají, že náročnost této úrovně je příliš vysoká (úroveň B2 pokládá jeden dotazovaný za obecně příliš vysokou; další dotazovaný nesouhlasí s požadavkem na užívání složitějších souvětí a abstraktního vyjadřování).

Pro úroveň *mírně pokročilý* souhlasili s *počtem dílčích testů* všichni dotazovaní, u časových limitů pro tuto úroveň se tři respondenti domnívali, že by bylo potřebné je zvýšit, jeden respondent nedokázal dostatečnost časového limitu posoudit. S *pocty bodů* pro jednotlivé dovednosti souhlasili všichni dotázaní. Co se týká úrovně *středně pokročilý*, většina respondentů souhlasí s *počtem dílčích testů*, pouze dva dotazovaní se domnívají, že užití jazyka není potřebné testovat odděleně. S *časovými limity* souhlasili všichni respondenti s výjimkou dvou – jeden nedokázal vhodnost časového limitu posoudit, druhý doporučuje časový limit pro čtení s porozuměním a psaní snížit, protože se žáci nedokáží tak dlouho soustředit. S *počtem bodů* pro jednotlivé dovednosti souhlasili všichni respondenti.

Všichni dotazovaní se domnívají, že maturitní zkouška v souladu s navrhovaným alternativním katalogem by měla na výuku anglického jazyka na středních školách *pozitivní dopad*. Důvody, které uvedli, byly různé, například:

- „Vedlo by to k tomu, že by studenti skutečně dokázali komunikovat a nejenom psát testy založené na gramatice a slovní zásobě.“
- „Bylo by přesně dané, co studenti na jednotlivých úrovních musí umět, takže i hodnocení by bylo objektivnější.“
- „Určitě by to ovlivnilo výuku směrem ke komunikativnímu přístupu.“
- „Srovnatelnost výsledků žáků různých škol a různých typů škol (přijímání na VŠ či do zaměstnání). Nový návrh řeší testování jazykových dovedností a ne např. reálií (dnes musí žák mít u maturity nejen jazykové, ale i faktografické znalosti...). Díky postavení požadavků na základech SERR může žák získat vysvědčení použitelné v rámci Evropy, dnešní maturita z jazyka příliš velkou výpovědní hodnotu nemá.“

Připomínky uvedené v dotazníku a v neformálních rozhovorech byly často velmi rozdílné. Zatímco někteří oslovení vítají skutečnost, že v alternativním katalogu se nevyskytuje požadavek na znalosti reálií a katalog se naopak koncentruje na vymezení požadavků na jazykové znalosti a řečové dovednosti, jiný dotazovaný si není jistý, jestli vyřazení zkoušení reálií je správné. Různí se také představy o obtížnosti zkoušky. Někteří respondenti se domnívají, že zkouška B2 je pro maturanty jen stěží dosažitelná, pro jiné se tato úroveň zdá být nízká, protože jí někteří žáci dosáhnou už v nižším než maturitním ročníku, což je vzhledem ke značným rozdílům mezi školami zcela pochopitelné. Jednomu z respondentů chybí přepočítání bodů na známky. Dva respondenti upozorňují, že na takto navrženou maturitní zkoušku je potřeba pedagogy středních škol vyškolit. Jeden respondent uvádí, že „ve srovnání s Maturitou nanečisto jsou srozumitelně dány potřeby, požadavky, kritéria i náplň. Nová maturita (CERMAT) se neodpoutala od překladu a stále zahrnuje i testování gramatiky (často výjimek), paměti a inteligence“.

### Dílčí závěr

Za nejdůležitější zjištění lze pokládat, že všem respondentům se jeví formulace požadavků srozumitelné. Různí se názory na časové limity, avšak tento problém lze vyřešit pouze tím, že se test před zadáním ověří. Zajímavé je, jak se liší názory na obtížnost jednotlivých úrovní maturity a na realističnost dosažení obou úrovní. Poukazuje to pouze na skutečnost, že existují velmi závažné rozdíly v úrovni žáků na různých školách stejného zaměření i na školách různých typů (gymnázia, střední odborné školy). Zavedení jednotné zkoušky je vesměs vnímáno jako pozitivní, avšak zároveň velmi komplikované z již uvedených důvodů. Názory se také různí na oddělené testování slovní zásoby a gramatiky u úrovně středně pokročilý a na testování reálií. Jak již bylo zmíněno, většina mezinárodně uznávaných zkoušek pro vyšší úroveň část týkající se používání jazyka obsahuje. Zkoušení reálií naopak nemá s ověřováním řečových dovedností nic společného. Středoškolské pedagogy se domnívají, že je sice důležité naučit žáky cizojazyčné komunikaci v každodenním životě, nicméně poukazují na skutečnost, že musí mnoho času věnovat i gramatice, protože u přijímacích zkoušek na vysoké školy je

toto více-méně jediný způsob ověřování jazykové úrovně uchazečů. Tuto skutečnost lze hodnotit jako negativní.

Všechny informace získané od učitelů anglického jazyka a rodičů poskytly cenné názory a připomínky.

## 10. Konstrukce testu podle alternativního katalogu požadavků

Podle připraveného alternativního katalogu požadavků k maturitní zkoušce byly pro úroveň mírně pokročilý navrženy čtyři subtesty zjišťující úroveň jednotlivých řečových dovedností. Všechny subtesty prošly standardním procesem připomínkování na mém pracovišti. Na připomínkování se podíleli dva kolegové s několikaletou zkušeností v oblasti jazykového testování. Testy nebyly předloženy k připomínkování středoškolským učitelům před administrací testu, a to z toho důvodu, že v převážné většině středoškolských učitelé nejsou pro tvorbu testů kvalifikovaně vyškoleni.

Následující dvě tabulky stručně shrnují nejdůležitější specifikace testu ověřujícího receptivní a produktivní dovednosti.

Tab. 6. Specifikační tabulka dílčích testů ověřujícího receptivní dovednosti

makro dovednosti	mikro dovednosti	čas. limit	položky (body)	testovací techniky	počet textů	vlastnosti nahrávek
poslech	rozpoznat jednoduchou slovní zásobu a fráze; identifikovat hlavní informace; identifikovat jednoduché detaily	20 až 25 minut	25	výběr z odpovědí; krátké odpovědi; doplňování; přiřazování; pravdivá – nepravdivá tvrzení	2 – 3 monology krátké až středně dlouhé; 2 – 3 dialogy krátké až středně dlouhé;	jednoduché; jasná výslovnost; témata každodenního života;
čtení	rozpoznat jednoduchou slovní zásobu a fráze; identifikovat hlavní informace; identifikovat jednoduché detaily	45 až 50 minut	25	výběr z odpovědí; krátké odpovědi; doplňování; přiřazování; pravdivá – nepravdivá tvrzení	4 – 6 krátkých textů; 4 – 6 středně dlouhých textů	jednoduché; témata každodenního života; autentické, minimálně upravované

U časových limitů pro subtesty obou dovedností je uvedeno rozpětí s rozdílem pěti minut (20 – 25 a 45 – 50), protože konečný časový limit je možné pro každý subtest stanovit pouze po pretestování na dostatečně velkém vzorku cílové populace. Přestože počet položek (počet bodů) je pro každý subtest uveden 25, navrhované ukázkové subtesty obsahují 30 položek. Tento vyšší počet je úmyslný, protože v procesu ověřování se některé položky mohou projevit jako nevhodné. Typy testovacích technik zahrnují seznam možných použitých technik. Počty textů použité pro testování čtení i poslechu nejsou také dány striktně, protože by to velmi limitovalo tvůrce testů. Důležité je, aby byl dodržen požadovaný počet testových položek a časové limity. Pro každou dovednost se doporučuje použít několik textů různého tématického zaměření a také několik různých testovacích technik.

Tab. 7. Specifikační tabulka testu ověřujícího produktivní dovednosti

makro dovednosti	časový limit	počet bodů	počet testových úloh	typy testových úloh
psaní	35 – 40 minut	25	3	1 krátký strukturovaný vzkaz 2 delší strukturované písemné projevy
mluvení	15 minut	25	3	1 monolog s doplňujícími otázkami na známé téma; 2 sehrání situační role; 3 získání a předání informací

Časový limit pro subtest dovednosti psaní je také dán rozsahem 35 – 40 minut, protože jeho vhodnost je potřeba také ověřit. Pro obě produktivní dovednosti jsou doporučené tři testové úlohy, což je plně v souladu s odbornou literaturou.

### 10.1 Poslech s porozuměním (viz příloha P - 4)

Pro zjišťování úrovně poslechu byl vytvořen dílčí test obsahující 30 testových položek. Byl navržen žákovský sešit, odpovědní arch, klíč s odpověďmi a přepis nahrávky. Tématické zaměření nahrávek, obtížnost použitého jazyka a typy testovacích technik byly vybrány v souladu s požadavky formulovaným v alternativním katalogu a se specifikační tabulkou (tab. 6). Tento dílčí test nebyl v praxi ověřován z důvodu složitější administrace (magnetofon, sluchátka, nejlépe jazyková laboratoř) a rozsahu této práce.

#### Tematické zaměření nahrávek

Témata nahrávek pokrývají běžné komunikační situace každodenního života, jak uvádí alternativní katalog (zanechání vzkazu na telefonickém záznamníku, budoucí povolání, pasová kontrola, životopisní data, rodina, trávení volného času, turistické informace). Nahrávky obsahují jak monology, tak i dialogy (viz tab. 6).

## **Testovací techniky**

Testovací techniky v tomto dílčím testu byly vybrány v souladu s doporučeními současné odborné literatury (viz ČÁST I). Použité techniky také korespondují s technikami uvedenými v alternativním modelu Katalogu požadavků (výběr odpovědi z volby a, b, c nebo d, krátké odpovědi na otázky, přenos informací z nahrávky, pravdivá a nepravdivá tvrzení, doplňování chybějících informací do shrnutí) a se specifikační tabulkou (tab. 6).

### **10.2 Čtení s porozuměním (viz příloha P - 5)**

Pro zjišťování úrovně čtení byl vytvořen dílčí test obsahující 30 testových položek. Tento dílčí test byl ověřován v praxi. Byl navržen žákovský sešit, odpovědní arch, klíč s odpověďmi, průvodní instrukce k zadávání testu a poté stručný dotazník zjišťující zpětnou vazbu od učitelů. Doba trvání byla stanovena na 45 minut a úkolem ověření testu bylo kromě jiného také zjistit, jestli je tento časový limit odpovídající.

### **Tematické zaměření textů**

Dílčí test hodnotící dosaženou úroveň čtení vycházel výhradně z autentických textů týkajících se každodenního života. V případě potřeby byly některé texty pouze nepatrně adaptovány, a proto se jejich obtížnost může na první pohled jevit trochu vyšší, než se požaduje. Z tohoto důvodu byla záměrně snížena obtížnost testových úloh. Test obsahuje 11 textů různého tematického zaměření v souladu s požadavky alternativního katalogu (volný čas, nákupy, parkování ve městě, ubytování na kolejích, program pro zahraniční studenty, zdravý životní styl, nemoci, turistické informace, informace o kurzech angličtiny, jak přestat kouřit, služby).

## **Testovací techniky**

Testovací techniky byly stanoveny podle doporučení alternativního katalogu, specifikační tabulky (tab. 6) a v souladu se současnými trendy v jazykovém testování. Bylo použito těchto technik: přiřazování, doplňování chybějících informací, pravdivá a nepravdivá tvrzení, výběr odpovědi z volby a, b, c nebo d, krátké odpovědi na otázky.

### **10.3 Psaní (viz příloha P - 6)**

Pro ověřování úrovně psaní byl navržen dílčí test, který obsahuje tři testové úlohy. První požaduje od kandidátů napsat jednoduchý vzkaz, druhá krátký dopis a třetí stručný email. Každá úloha byla zasazena do určitého kontextu a pro každou úlohu byla připravena kritéria hodnocení. I tento dílčí test byl ověřován v praxi. Byl navržen žákovský sešit, odpovědní arch, průvodní instrukce k zadávání testu a poté stručný dotazník zjišťující zpětnou vazbu od učitelů. Doba trvání byla stanovena na 35 minut a úkolem ověření testu bylo kromě jiného také zjistit, jestli je tento časový limit odpovídající.

### **Tematické zaměření**

Tematicky byly testové úlohy zaměřeny takto: první se týkala komunikace kandidáta s hostitelskou rodinou v anglicky mluvící zemi (napsání vzkazu), účelem druhé úlohy bylo podání základních informací o kandidátovi (vzdělání, zájmy, rodina), zdůvodnění, proč se chce věnovat práci au-pair v zahraničí a popsání vlastních zkušeností s péčí o děti (krátký dopis). Třetí testová úloha vyžadovala napsání emailu zahraničnímu příteli, ve kterém by kandidát příteli doporučil, jaká místa v České republice navštívit, kde se ubytovat, jak cestovat apod.

### **Hodnotící kritéria (viz příloha P - 8)**

Pro hodnocení písemného projevu byla zformulována hodnotící kritéria, a to pro každou testovou úlohu zvlášť. Kritéria vycházela zejména z požadavků na úroveň mírně pokročilý (A2), což v praxi obecně znamená, že kandidáti se vyjadřují jednoduše, s chybami, ale tak, že by jim rodilý mluvčí nebo nerodilý mluvčí komunikující anglicky rozuměl.

### **10.4 Mluvení (viz příloha P -7)**

Test zjišťující úroveň dovednosti mluvení nebyl v praxi z časových důvodů ověřován. Pro testování dovednosti mluvení byl přesto zpracován popis zkoušky obsahující také charakteristiku tří testových úloh. Kromě toho byla vypracována i hodnotící kritéria a tři ukázkové testové úlohy.

### **Tematické zaměření**

Testové úlohy zjišťující úroveň dovednosti mluvení se podobně jako ostatní testové úlohy řídí požadavky alternativního katalogu. Tematické zaměření je v souladu s alternativním katalogem a týká se zejména situací každodenního života.

### **Typy testových úloh**

V souladu s odbornou literaturou se v dílčím testu zjišťujícím úroveň mluvení používají tři různé typy testových úloh. První testová úloha má formu krátkého rozhovoru (*interview*), kdy se kandidát stručně představí a examinátoři se ptají další doplňující otázky. Druhým typem testové úlohy je sehrání situační role (*role play*) v běžném kontextu každodenního života. Ve třetí testové úloze má kandidát zjistit od examinátora určité informace (*information gathering task*).

### **Hodnotící kritéria (viz příloha P - 9)**

Byla vypracována hodnotící kritéria, která popisují požadovaný ústní projev pro přidělení určitého počtu bodů za jednotlivé testové úlohy. Důraz je kladen na splnění komunikačního záměru, správné používání jazyka a specifikaci přípustných chyb.



## 11. Zadání a hodnocení subtestů

### 11.1 Zadání subtestů

Z časových důvodů a vzhledem k rozsahu této práce byly v praxi ověřovány pouze dva ze čtyř navržených subtestů. Jednalo se o subtesty zjišťující úroveň čtení a psaní. Oba subtesty nebo pouze jeden z nich byly zadány celkově 126 žákům a studentům. Oba subtesty napsalo 85 žáků a studentů. Důvodem, proč nebyly oba subtesty zadány stejným kandidátům, bylo množství času, které byli učitelé ochotni psaní testu věnovat. Dalším důvodem bylo, že protože se subtesty nepsaly v jedné hodině, někteří žáci a studenti mohli v jedné ze dvou hodin chybět a tudíž napsali pouze jeden ze dvou subtestů.

Kandidáti, kteří psali subtesty, byli z různých škol. Záměrně nebyli vybíráni pouze z maturitních ročníků, protože jak již bylo zmíněno, mezi různými školami existují v dosažené jazykové úrovni velké rozdíly. Za účelem vybrání vhodných kandidátů byli osloveni učitelé středních a jedné vysoké školy s požadavkem, aby vybrali skupinu nebo skupiny, jejichž úroveň odpovídá úrovni „mírně pokročilý“, tedy přibližně úrovni A2 podle SERR. Toto kritérium se jeví být pro výběr odpovídajících kandidátů vhodnější než jejich automatický výběr z maturitních ročníků.

Testy byly administrovány na Fakultě vojenských technologií University obrany (FVT UO, Brno), Gymnáziu Křenová (Brno), Gymnáziu Slovanské náměstí (Brno), Vojenské střední škole (Moravská Třebová), Gymnáziu Vídeňská (Brno) a na Gymnáziu v Tišnově (Tišnov) v období listopad 2005 – únor 2006. Obecně se učitelé středních škol stavěli k zadávání testu váhavě, nedůvěřivě až neochotně.

### 11.2 Hodnocení testů

Tato subkapitola se zabývá způsoby hodnocení obou subtestů, z nichž jeden byl hodnocen většinou objektivně (subtest čtení), zatímco druhý (subtest psaní) byl hodnocen subjektivně. Jsou uvedeny připomínky k subtestům zejména ze strany učitelů, ale také ze strany žáků a stanovisko autorky k nim. Autorka také shrnuje své postřehy získané během opravy testů.

#### 11.2.1 Subtest zjišťující úroveň čtení

Tento subtest byl hodnocen výlučně autorkou podle předem připraveného klíče. Učitelé středních škol ve většině případů byli nanejvýš ochotni test zadat, ne však ho už opravovat. Tuto skutečnost bylo nutno akceptovat, přestože zapojení více hodnotitelů by mohlo objevit možné nejasnosti či obtíže při hodnocení. Položky obsažené v testu byly většinou hodnoceny objektivně (přirazování -10, pravdivá a nepravdivá tvrzení - 5, výběr odpovědi z vícenásobné nabídky - 3; celkem 18 položek). Ze zbývajících dvanácti položek bylo devět doplňovacích a tři položky představovaly otázky s krátkými odpověďmi.

Dotazník k tomuto subtestu (viz příloha P - 10) zjišťoval základní fakta o testu - jestli byl časový limit dostačující a jestli byly instrukce srozumitelné. Dále pak byl dán prostor připomínkám zejména ze strany učitelů. Pokud však existovaly připomínky i ze strany žáků nebo studentů, učitelé mohli i tyto připomínky do dotazníku zaznamenat.

### **Připomínky a doporučení ze strany učitelů zadávajících dílčí test**

Z devíti dotázaných učitelů pět uvedlo, že časový limit 45 minut pro tento dílčí test byl nedostatečný, podle čtyř zbývajících učitelů byla časová dotace dostačující. Osm z devíti dotázaných pokládá instrukce za srozumitelné, jeden respondent uvádí několik připomínek zahrnujících i připomínky žáků: „chybí instrukce jak opravit v případě chybného záznamu; zda zatrhnout vždy pouze jednu volbu; jestli při odpovědích na otázky je potřeba odpovědět celými větami nebo zda stačí jednoslovná odpověď.“

Jedna skupina studentů FVT UO označila test jako středně těžký. Jiná skupina studentů FVT UO pokládá texty za zajímavé. Podle jednoho středoškolského učitele „řada výrazů neodpovídá slovní zásobě střední školy na daném stupni, např.: *in-stock selection, merchandise, toll-free, shipping, inventory sheet, quarry*.“ Podobného názoru je další vyučující, která se domnívá, že na úroveň A2 je slovní zásoba textů příliš složitá a doplňuje, že soudí tak podle používaných učebnic. Další připomínka se vyskytovala u dvou respondentů a poukazovala na to, že v odpovědním archu je málo místa na odpovědi. Podle jedné skupiny studentů jim chyběl čas na zkontrolování, a navrhuje časový limit zvýšit na 60 minut.

### **Postřehy z procesu opravy testu**

Jak už bylo zmíněno, testové úlohy i navržené správné odpovědi byly konzultovány a připomínkovány v užším pracovním týmu. Všechny relevantní připomínky byly zapracovány. Přesto však fáze ověřování testu, respektive proces opravování testových položek byl velmi přínosný, protože u krátkých odpovědí byly v odpovědích žáků nalezeny i některé další možnosti, než které uváděl předem připravený klíč. V převážné většině však bylo jednoduché rozhodnout, jestli tyto odpovědi hodnotit jako správné nebo ne. Dále je uveden seznam případů, ve kterých se odpovědi žáků mírně lišily od očekávané odpovědi (klíče), ale byly akceptovány, protože signalizovaly, že žáci textu porozuměli.

- Pro položku 4 uvádí klíč správnou odpověď *two*; při opravování testu se objevila také odpověď *different* nebo *two different*. Obě možnosti byly akceptovány.
- V položce 5 je podle klíče správná odpověď *practise* (sloveso); za správnou byla však také pokládána odpověď *practice* (podstatné jméno), kromě toho byla akceptována také odpověď *understand*, protože je smysluplná.
- V položce 22 byla jako správná odpověď uvedena tato: *Tell everyone you are stopping*. Další významově možné odpovědi byly také akceptovány (*No thanks, I don't smoke; chewing gum; stop-smoking groups; hypnosis or acupuncture therapy; say I don't smoke*).

- Klíč uvádí pro položku 25 jako správnou odpověď *trained/qualified*, ale byly akceptovány i další významově odpovídající odpovědi (*professional, helpfull* (chybný pravopis), *pleased to offer help*).
- U položky 26 vedle klíčem stanovené správné odpovědi *advice*, byly akceptovány i odpovědi *recomanding* (chybný pravopis) a *advices* (nesprávný tvar).
- Podobné situace nastaly i u položek 27 - 30, u kterých byly také akceptovány odpovědi významem odpovídající.

### Stanovisko k připomínkám a doporučením

Fáze ověřování testů na odpovídající populaci je nedílnou a velmi důležitou částí testovacího procesu. Jak ukazuje i ověřování tohoto dílčího testu čtení, pouze odpovědi kandidátů a následná analýza testu může podat objektivní informaci o tom, jestli test jako takový a také jednotlivé jeho položky jsou správně navrženy.

Připomínky ze strany učitelů i žáků jsou velmi cenné. Zároveň je potřeba uvědomit si, že učitelé vyjadřující se k testu mají sice odpovídající vzdělání a zkušenosti ve výuce anglického jazyka, avšak nejsou školeni v oblasti jazykového testování. Za relevantní je možné pokládat potřebu prodloužit časový limit pro vypracování testu (doporučuje pět z devíti dotázaných učitelů); dále pak doplnit a upřesnit instrukce k některým úlohám testu. V úvodu testu by měl být uveden doporučený způsob opravy špatně označené odpovědi. U krátkých odpovědí by pak bylo vhodné doporučit, že stačí odpovědět stručně, několika slovy, ne celou větou. Připomínka týkající se nedostatku prostoru pro odpovědi je také relevantní – test by měl poskytnout více místa pro odpověď na odpovědních arších.

Nelze však souhlasit s námitkami týkajícími se neadekvátní slovní zásoby použité v textech. Texty byly záměrně vybrány z autentických materiálů a pouze minimálně adaptovány, protože jsou to typy textů, se kterými se žáci v cizojazyčném prostředí setkávají. V reálném životě jim nikdo nebude existující texty zjednodušovat. Navíc vzhledem k tomu, že texty jsou autentické, byly testové úlohy navrženy záměrně jednodušší. V praxi to znamená, že ke správné odpovědi nemusejí kandidáti znát všechnu slovní zásobu, která se v textu nachází. Stanovisko učitelů k této problematice je nesprávné; tento druh zkoušky nemá za cíl ověřit zvládnutí určitého kvanta slovní zásoby podle určité učebnice; naopak, má za cíl zjistit, jestli kandidáti dosahují úrovně, které specifikuje alternativní katalog požadavků.

#### 11.2.2 Subtest hodnotící úroveň psaní

Původním záměrem bylo, aby podle kritérií hodnocení ohodnotili subtest psaní jednak učitelé, kteří test zadávali, jednak nezávisle na nich autorka. Učitelé středních škol však až na jednu výjimku odmítli testy hodnotit z důvodu vysoké pracovní vytíženosti. Aby se zjistilo, jestli jsou kritéria hodnocení funkční, bylo přesto potřeba, aby testy psaní ohodnotili dva na sobě nezávislí hodnotitelé. Proto byli na bázi dobrovolnosti do hodnocení zapojeni učitelé z pracoviště autorky (katedra jazyků FVT UO). Každý

hodnotitel měl k dispozici hodnotící kritéria písemného projevu (příloha P - 8) a hodnocení se zapisovala na zvláštní papír (ne do testu, aby se hodnotitelé vzájemně neovlivňovali).

Dotazník k tomuto dílčímu testu se zaměřoval na zjištění, jestli byl časový limit dostačující, jestli byly instrukce srozumitelné a jestli bylo pro hodnotitele možné bez větších obtíží hodnotit testové úlohy psaní s použitím kritérií hodnocení (viz příloha P - 11). I v tomto dotazníku byl dán prostor pro připomínky učitelů, případně i žáků nebo studentů.

### **Připomínky a doporučení ze strany učitelů zadávajících subtest**

Z devíti dotázaných učitelů, kteří test zadávali, šest uvádí, že časový limit byl dostačující. Tři se domnívají, že časový limit byl příliš krátký. Všechny devět učitelů také pokládá instrukce byly srozumitelné. Kromě autorky vykonávalo hodnocení celkem sedm hodnotitelů a všichni uvedli, že hodnotit subtest psaní podle poskytnutých kritérií bylo pro ně jednoduché.

Jedna učitelka se domnívá, že úkoly byly zajímavé a odrážejí situace, do kterých by se žáci mohli dostat. Jiná vyučující, přestože odpověděla kladně na otázku, jestli pro ni bylo jednoduché hodnotit psaní podle kritérií, uvádí, že se „nedokáže 100% rozhodnout, zda její míra porozumění by byla stejná u rodilého mluvčího“.

Jiná respondentka se v části dotazníku vyhrazené pro připomínky ptá, jestli se žáci mohou dotazovat učitele, který test zadává, a to zejména na zadání a případně i jiné výrazy. Sama navrhuje zadání v češtině a doporučuje „definovat rozsah počtem slov.“ Z připomínek ze strany žáků zmiňuje, že shledávají úkol obtížný a myslí si, že by mohli používat slovník a že by potřebovali více času na kontrolu.

Další dotázaná doporučuje, aby bylo uvedeno, jestli má být dopis (druhá testová úloha) napsán formou dopisu, jak mají být jednotlivé testové úlohy dlouhé a jestli požadované informace mají být uváděny v tom pořadí, jak jsou uvedeny v zadání.

Na jednom z gymnázií (Slovanské nám. Brno) byl test zadáván ve dvou třídách a protože obě byly označeny jako třetí ročník, zaujalo mne, jak na první pohled rozdílných výsledků dosáhly (III.A. dosáhla lepších výsledků než 3.C). Cituji vysvětlení jedné z vyučujících, která mi ho zaslala emailem. „III. A je třetí ročník šestiletého cyklu s rozšířenou výukou angličtiny, tedy prváci věkem. 3.C je třetí ročník čtyřletého cyklu všeobecného zaměření, který s angličtinou začínal v prvním ročníku. Zde máte příklad, jak je takřka nemožný úkol sjednotit maturity a osnovy! Už vůbec si nedovedu představit odborné školy a učiliště.“

### **Postřehy z procesu opravy subtestu**

Opravování subtestu psaní bylo mnohem náročnější na čas i na samotné hodnocení než hodnocení subtestu čtení. Každý jednotlivec byl vždy hodnocen nezávisle dvěma

hodnotiteli a pouze v případech (ojedinělých), kdy se jejich hodnocení lišilo o více než dva body, se hodnotitelé sešli a s pomocí kritérií prošli testy ještě jednou a nakonec vždy došli ke kompromisnímu řešení, se kterým oba souhlasili.

V průběhu opravování se vyskytly určité problémy, respektive vyvstaly otázky, které by v případě dalšího používání testu „naostro“ bylo potřebné vyřešit. Každá ze tří testových úloh byla určena určité osobě a tudíž se očekávalo, že kandidáti při vypracovávání otázek na začátku osloví příjemce. V mnohých případech ale jakékoliv oslovení chybělo. Pro další používání testu by proto bylo vhodné zabývat se tímto problémem – upozornit kandidáty, že v testových úlohách je potřeba použít odpovídající formu (vzkaz, dopis, email) a že její použití bude součástí hodnocení. Tuto změnu by pak bylo nutné zahrnout do hodnotících kritérií.

V úloze 1 bylo v několika případech z odpovědí kandidátů zjevné, že instrukce nepochopili. Navíc často nezahrnuli do psaní obsahovou položku 3 *reason* (důvod, proč odcházejí do města). Je možné, že toto slovíčko neznají.

U této úlohy existují také velké rozdíly v úrovni jednotlivých žáků; zatímco někteří kandidáti mají v psaném projevu chyby a je jasné, že úkol plní s námahou, jiní ho naopak plní s lehkostí a do role se lehce vžívají. Při hodnocení bylo velmi důležité mít neustále na paměti, že body se nepřidělují tak, aby odrážely rozdíly mezi lepšími a horšími kandidáty. Naopak, stejný počet bodů dosáhli všichni kandidáti, kteří splnili požadovaná kritéria. Za zmínku nepochybně stojí i zjištění, že kandidáti, kteří byli na vyšší úrovni, nabízeli v této úloze často pro bod *no food in the fridge* i řešení situace, např. „*There wasn't any food in the fridge, so I bought some.*”

U testové úlohy 2 bylo zajímavé sledovat obsahově podobně zaměřené odpovědi různých skupin kandidátů. Studenti Univerzity obrany například jako důvod, proč chtějí pracovat jako *au pair*, nejčastěji uvádějí kromě zdokonalení se v angličtině možnost vydělat si peníze (většinou hoši); naproti tomu žáčky Gymnázia Křenová kromě touhy zdokonalit se v angličtině většinou uvádějí skutečnost, že mají rády děti, chtějí poznávat jinou zemi a nové lidi. V této úloze kandidáti často neuvedli vhodné informace o své rodině; pro uvedený účel (při poskytnutí informací pro rodinu, ve které mají kandidáti pracovat jako *au pair*), není relevantní, resp. dostačující uvádět jména rodičů, ale je vhodnější uvést jejich zaměstnání, své záliby, rodinné vztahy, atmosféru v rodině apod.

Zejména u úkolů 1 a 3 se mnozí kandidáti snažili do odpovědí zahrnout své znalosti reálií (britských a českých v angličtině) – často zmiňovali např. Charles Bridge, Wenceslav Square, Stonehenge.

## Stanovisko k připomínkám a doporučením

Podle vyjádření učitelů, kteří tento test zadávali, se časový limit jeví být odpovídající. Instrukce až na několik výjimek byly hodnoceny jako srozumitelné. Nelze souhlasit s doporučením jedné z vyučujících použít instrukce v češtině, protože tím by se test umístil do nepřirozeného prostředí. Pokud se pro instrukce používá jednoduchý jazyk, neměl by vzniknout problém nepochopení instrukcím. Navíc pokud by žáci chtěli vykonat jakoukoliv standardizovanou, mezinárodně uznávanou zkoušku z anglického jazyka, setkali by se pouze s instrukcemi v anglickém jazyce. Navíc lze předpokládat, že hodiny anglického jazyka jsou na středních školách vedeny alespoň částečně v angličtině a jazyk zadávání instrukcí ke cvičením v hodinách angličtiny je velmi blízký jazyku instrukcí k testovým úlohám v testu. Proto by ani nebylo vhodné, aby žáci měli mít možnost dotazovat se učitele na vysvětlení instrukcí případně některých výrazů v češtině. Přesto bývá běžnou praxí, že je provedena instruktáž v rodném jazyce, aby se eliminovaly problémy vzniklé s nepochopením instrukcím. Z tohoto důvodu by v případě dalšího používání tohoto testu bylo vhodné poskytnout kandidátům vysvětlení instrukcí v češtině a v zadání testu ponechat instrukce v angličtině.

Protože u některých studentů vznikl v testové úloze 1 pravděpodobně problém s porozuměním slovíčku *reason* (důvod), bylo by vhodné nahradit toto slovo slovem *why* (proč), které by mělo být všem žákům známé.

## 12. Interpretace výsledků subtestů

### 12.1 Subtest čtení

Tento subtest byl zadán celkově 110 kandidátům (většinou žákům středních škol, ale také studentům vysoké školy). Obsahuje 30 položek, které byly hodnoceny dichotomicky, tedy buď jako správné (1) nebo nesprávné (0). Po opravě subtestu byly výsledky vyhodnoceny softwarovým programem ITEMAN, který zpracoval položkovou analýzu (viz příloha P - 12). Položková analýza „má přispět k odhalení úloh, které nevyhovují účelu testu a úloh s technickými nedostatky“ (Byčkovský, Zvára, s. 29).

Nejnižší počet bodů (8) dosáhl jeden kandidát, nejvyššího počtu bodů (30) dosáhli tři kandidáti. Průměrný hrubý skóre (*mean*) u testu byl 20,73. Standardní odchylka pro tento test se rovná 5,40.

**Reliabilitu** (*reliability*) testových výsledků lze ověřovat různými způsoby. V našem případě analýzy testových výsledků ji program ITEMAN vyjadřuje pomocí Cronbachova koeficientu alfa. Ten je vypočten na základě korelace jednotlivých položek s celkovým skórem. Obecně lze říci, že „výpočet reliability umožňuje (...) posoudit, jak dalece jsou výsledky testování ovlivněny náhodou a jak jsou tudíž věrohodné“ (Chráska, 1999, s. 58). Jinými slovy, hodnota alfa vyjadřuje v této analýze testu koeficient spolehlivosti a je indexem homogenity položek v rámci daného testu. Může nabývat hodnot v intervalu 0,00 – 1,00. (User's Manual for the ITEMAN, 1996, s. 3 - 17). U našich testových výsledků se rovná 0,844. Takovýto výsledek lze pokládat za uspokojivý. Uvedené

statistické charakteristiky testu přehledně prezentuje obr. 21 (výstup softwarového programu ITEMAN). Překlad jednotlivých charakteristik do češtiny byl přebrán z prostudované literatury (Byčkovský, Zvára, s. 39)

Scale Statistics		
-----		
Scale:	1	
	-----	
N of Items	počet položek	30
N of Examinees	počet testovaných	110
Mean	průměrný hrubý skór	20.727
Variance	rozptyl skórů	29.198
Std. Dev.	směrodatná odchylka skórů	5.404
Skew	šikmost rozložení skórů	-0.137
Kurtosis	špičatost rozložení	-0.975
Minimum	nejnižší dosažený skór	8.000
Maximum	nejvyšší dosažený skór	30.000
Median	prostřední skór	20.000
Alpha	reliabilita (Cronbachovo alfa)	0.844
SEM	směrodatná chyba měření	2.133
Mean P	relativní průměrný skór	0.691
Mean Item-Tot.	průměrný bodově biseriální koeficient	0.418
Mean Biserial	průměrný biseriální koeficient	0.593
Max Score (Low)	nejvyšší skór dosažený skupinou 27% testovaných s nejnižšími skóry (horší skupina)	16
N (Low Group)	počet testovaných v horší skupině	31
Min Score (High)	minimální skór dosažený skupinou 27% testovaných s nejvyššími skóry (lepší skupina)	25
N (High Group)	počet testovaných v lepší skupině	33

Obr. 21. Statistické charakteristiky testu

**Rozložení** (*distribution*) může obecně být normální ve tvaru zvonu (*normal, bell-shaped*), zkosené (*skewed*) reprezentované asymetrickou křivkou, která signalizuje, že většina skórů se nachází na jednom konci křivky, nebo bimodální (*bimodal*), u kterého se skóry hromadí ve dvou rozdílných bodech (Alderson; Clapham; Wall 1995, s. 94-95). Rozložení u analyzovaného testu je znázorněno na obr. 22 a je bimodální, což je považováno za ideální právě pro testy, jejichž účelem není uspořádat výsledky kandidátů podle pořadí, ale rozhodnout, jestli kandidáti požadovanou úroveň (skóre) dosahují nebo ne. Z obrázku 22 lze usoudit, že za hraniční skóre je možné pokládat hodnotu kolem 21 bodů.

SCALE # 1

## Score Distribution Table

Number Correct	Freq- uency	Cum Freq	PR	PCT	
. . . No examinees below this score . . .					
7	0	0	1	0	
8	1	1	1	1	#
9	1	2	2	1	#
10	0	2	2	0	+
11	2	4	4	2	##
12	1	5	5	1	#
13	7	12	11	6	#####
14	3	15	14	3	###
15	6	21	19	5	+#####
16	10	31	28	9	#####
17	3	34	31	3	###
18	10	44	40	9	#####
19	3	47	43	3	###
20	8	55	50	7	+#####
21	2	57	52	2	##
22	6	63	57	5	#####
23	5	68	62	5	#####
24	9	77	70	8	#####
25	6	83	75	5	+#####
26	10	93	85	9	#####
27	5	98	89	5	#####
28	3	101	92	3	###
29	6	107	97	5	#####
30	3	110	99	3	+###

|-----+-----+-----+-----+  
5 10 15 20 25  
Percentage of Examinees

Obr. 22. Rozložení počtu správných odpovědí subtestu čtení.

**Obtížnost** jednotlivých položek je jednou z jejich základních charakteristik. Vyjadřuje, do jaké míry jsou kandidáti schopni daný test nebo testovou položku vypracovat správně. (Davies et al., 1999, s. 45). Obrázek 23 zobrazuje všechny hodnocené parametry dvou problematických položek (položka č. 9 a 21) s nejnižší obtížností (*facility value, item difficulty*). Na položku č. 9 odpovědělo správně 98% kandidátů, na položku č. 21 také vysoké procento kandidátů (95%). V pravé části počítačového výstupu jsou tučně vyznačeny také distraktory, které jsou zjevně nefunkční, protože je žádný kandidát nepoužil. Správné odpovědi jsou označeny hvězdičkou. Vzhledem k tomu, že tak nízkou obtížnost měly pouze dvě položky, nejde o závažný problém.



Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.
9	1-9	.98	.06	.22	A	.00	.00	.00	
					B	.00	.00	.00	
					C	.02	.06	.00	-.22
					D	.00	.00	.00	
					E	.98	.94	1.00	.22 *
					Other	.00	.00	.00	
21	1-21	.95	.06	.17	A	.00	.00	.00	
					B	.00	.00	.00	
					C	.95	.94	1.00	.17 *
					D	.04	.03	.00	-.13
					Other	.01	.00	.00	-.12

Obr. 23. Položková analýza dvou nejméně obtížných testových úloh

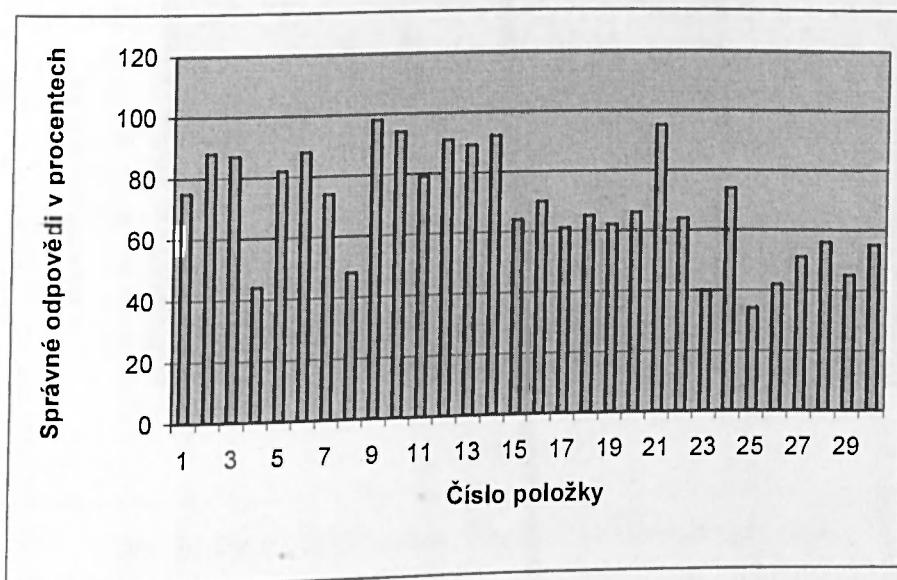
Naopak nejtěžšími položkami se ukazují položky č. 23 (0,40) a č. 25 (0,34). Charakteristiky těchto položek ukazuje obr. 24. V tomto případě se jednalo o testové úlohy s krátkou odpovědí a přibližně stejný počet kandidátů odpověděl na obě otázky správně i nesprávně.

Item Statistics					Alternative Statistics				
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.
23	1-23	.40	.75	.61	1	.40	.10	.85	.61 *
					2	.39	.45	.15	-.23
					Other	.21	.00	.00	-.46
25	1-25	.34	.79	.66	1	.34	.03	.82	.66 *
					2	.35	.23	.18	-.02
					Other	.31	.00	.00	-.66

Obr. 24. Položková analýza dvou nejvíce obtížných testových úloh

Neexistuje obecně platné pravidlo, které by předepisovalo nebo doporučovalo obtížnost položek. Většinou platí, že pokud je účelem testu co nejširší rozložení počtu bodů, doporučuje se vybírat položky s obtížností kolem 50% (Alderson; Clapham; Wall 1995, s. 81). Pokud však test má rozdělit kandidáty na ty, kteří uspěli a ty, kteří neuspěli, měla by se obtížnost pohybovat kolem hodnoty, která je požadovaná pro složení testu, např. pokud ke složení testu je potřeba dosáhnout 70% celkového počtu bodů, měla by se obtížnost položek pohybovat kolem 70% (Davies et al., 1999, s. 95 - 96). Ověřovaný test lze zařadit právě do této kategorie, protože jeho cílem je rozřadit kandidáty na úspěšné a neúspěšné. Byčkovský doporučuje k běžnému používání položky s obtížností 0,3 – 0,7 (Byčkovský, Zvára, s. 30).

Dle obrázku 25 lze říci, že většina položek se pohybuje mezi 60 – 80% úspěšných odpovědí. Chráska doporučuje, aby v testu nebylo příliš mnoho ani velmi obtížných ani velmi snadných úloh. Úlohy, které téměř žádný kandidát nevyřeší správně, doporučuje z testu odstranit, zatímco příliš snadnou úlohu „je možno z psychologických důvodů doporučit jako úvodní úlohu v testu.“ (Chráska, 1999, s. 47). Z uvedeného vyplývá, že ověřovaný test tento požadavek splňuje.



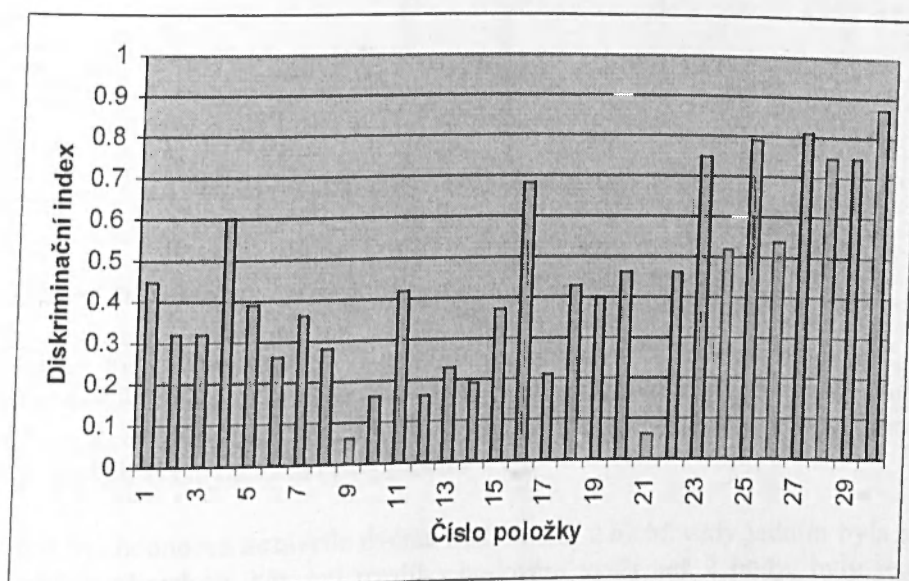
Obr. 25. Obtížnost jednotlivých položek subtestu čtení vyjádřena v procentech správných odpovědí

**Diskriminační index, citlivost úlohy (*discrimination index*)  $d$**  vyjadřuje, jak dobře určitá položka rozlišuje mezi lepšími a horšími kandidáty (*masters and non-masters*). Byčkovský pokládá citlivost úlohy za její důležitou charakteristiku a uvádí, že k plnění své základní funkce, „tj. měření, jak si studenti osvojili učivo nebo jaké mají předpoklady ke studiu..., by měly výsledky v testu dostatečně rozlišovat mezi lepšími a horšími studenty“ (Byčkovský, Zvára, s. 30). Ideální položka je taková, kterou všichni lepší kandidáti vypracují správně a všichni horší nesprávně. Neexistuje žádné pravidlo, které by stanovilo požadovanou hodnotu diskriminačního indexu. Je však zřejmé, že

diskriminační index by neměl nabýt záporné hodnoty: „Je žádoucí, aby každá z úloh měla  $d$  s co nejvyšší kladnou hodnotou. Úlohy, jejichž  $d$  má zápornou hodnotu, se považují za nevhodné“ (Byčkovský, Zvára, s. 30). V praxi by to totiž znamenalo, že určitou položku vypracovalo správně více horších než lepších studentů. Takovéto položky je nutno před použitím v „ostrém“ testu buď odstranit, nebo přepracovat a opět pretestovat.

V analyzovaném testu byl problematický (velmi nízký) diskriminační index u dvou položek (0,06). Jedná se o stejné položky, které byly zmíněny jako nejméně obtížné (položka č. 9 a 21) – viz. obr. 23 a 25.

Obr. 26 ukazuje, že u ostatních položek byl zjištěn diskriminační index od 0,16 do 0,87. Za pozitivní lze označit skutečnost, že žádná z položek neprokázala negativní hodnotu diskriminačního indexu. Pokud by se hodnocený test měl dále používat, bylo by vhodné položky č. 9 a 21 přepracovat a následně pretestovat.



Obr. 26. Diskriminační index u jednotlivých položek subtestu čtení

Za položky s nejvyšším diskriminačním indexem lze označit položky č. 25, 27 a 30. Jejich hodnoty uvádí obrázek 27. Je zajímavé, že položka č. 25 je zároveň i nejtěžší položkou testu (obr. 24). Ani dvě další položky, které mají vysoký diskriminační index, nepatří k položkám lehkým (51 a 55%). Uvedené úlohy s nejvyšším diskriminačním indexem byly úlohy s tvořenou odpovědí, ve kterých bylo také hodně odpovědí vynecháno.

Item Statistics				Alternative Statistics					
Seq. No. Key	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	High	Point Biser.
25	1-25	.34	.79	.66	1	.34	.03	.82	.66 *
					2	.35	.23	.18	-.02
					Other	.31	.00	.00	-.66
27	1-27	.51	.81	.67	1	.51	.06	.88	.67 *
					2	.16	.19	.12	-.09
					Other	.33	.00	.00	-.64
30	1-30	.55	.87	.70	1	.55	.13	1.00	.70
					2	.11	.06	.00	-.07 *
					Other	.35	.00	.00	-.69

Obr. 27. Položky s nejvyšším diskriminačním indexem

## 12.2 Subtest psaní

Tento subtest byl celkově zadán 101 kandidátům (většinou žákům středních škol, ale také studentům vysoké školy). Obsahoval celkově tři různé testové úlohy. U každé z nich byly v zadání uvedeny informace, které bylo potřeba zmínit. Ke každé testové úloze byla vypracována kritéria hodnocení (viz příloha P - 8).

Každý test byl hodnocen nezávisle dvěma hodnotiteli, z nichž vždy jedním byla autorka. V ojedinělých případech, kdy byl rozdíl v bodování vyšší než 2 body, byly jednotlivé testy hodnotiteli ještě jednou přečteny, výsledky poté konzultovány a bylo odsouhlaseno konečné hodnocení.

Průměr rozdílů v hodnocení všech subtestů psaní je 1,1 a korelace hodnocení autorky s ostatními hodnotiteli je 0,96. Tento výsledek lze pokládat za velmi úspěšný, protože naznačuje, že kritéria hodnocení jsou jasná a pomáhají různým hodnotitelům hodnotit velmi podobně. Nutno připomenout, že v případě „ostrého“ používání testu a tedy i hodnotících kritérií by bylo nezbytné připravit praktické školení pro všechny hodnotitele, čímž by měla být zajištěna ještě vyšší konzistentnost hodnocení.

## 12.3 Subtesty čtení a psaní

Oba subtesty psalo 85 žáků nebo studentů. Aby se potvrdilo, že každý subtest skutečně testuje jinou dovednost a že je proto potřebné zadávat oba subtesty, byl také vypočten

Pearsonův koeficient korelace mezi výsledky kandidátů v subtestu čtení s porozuměním a v subtestu psaní. Jeho hodnota 0,47 ukazuje, že výsledky u obou subtestů nekorelují vysoce, což potvrzuje, že každý subtest ověřuje jiné dovednosti a tudíž je vhodné používat oba subtesty.

### 13. Závěr

ČÁST I této práce se zabývala některými teoretickými otázkami souvisejícími s testováním. Jedná se o problematiku vzdělávacích cílů, problematiku tvorby katalogů požadavků a hodnotících nástrojů a o problematiku jazykového testování. U vzdělávacích cílů jsou zmíněny některé teoretické pohledy na formulování vzdělávacích cílů - zejména nejdůležitější myšlenky Magera a Blooma. Jsou vysvětleny nejdůležitější vlastnosti, kterými by hodnotící nástroje podle nich vytvořené měly disponovat - validita, reliabilita, objektivita a autenticita. Protože práce jako celek se zaměřuje na hodnocení studijních výsledků v anglickém jazyce, jsou také zmíněny a stručně charakterizovány nejnámější mezinárodně uznávané zkoušky z anglického jazyka.

Dále se tato část zabývá receptivními i produktivními řečovými dovednostmi, jejich definicemi a charakteristikou jejich společných i specifických rysů. Testovacím technikám jednotlivých řečových dovedností jsou věnovány samostatné oddíly. Jsou srovnány pohledy současných specialistů na jejich výhody i nevýhody. Závěrečná kapitola se vrací zpět k formulaci vzdělávacích cílů a kromě stručného popisu Bloomovy taxonomie výukových cílů prezentuje pokus zařadit jednotlivé testovací techniky do kategorií této taxonomie a tím aplikovat Bloomovu taxonomii na ověřování úrovně cizích jazyků.

ČÁST II zmiňuje nejzajímavější závěry průzkumu týkajícího se vedení maturitní zkoušky z anglického jazyka, kterým se autorka zabývala v roce 2002 a jehož cílem bylo srovnat obsah a formu maturitní zkoušky z anglického jazyka. Do průzkumu bylo zahrnuto šest středních škol v Brně (3 gymnázia a 3 střední odborné školy). Výsledky průzkumu jasně poukázaly na velké rozdíly mezi chápáním a samotným prováděním maturitní zkoušky na jednotlivých školách a na potřebu vytvořit podmínky pro zavedení nové koncepce.

Tato část dále obsahuje hodnocení původního Katalogu požadavků k maturitní zkoušce, ve kterém se podrobně rozebírají klady a nedostatky tohoto dokumentu. Samotný vznik Katalogu požadavků je hodnocen jako vysoce pozitivní, avšak jsou uvedeny i jeho nedostatky. Za nejzávažnější z nich lze označit například ne úplně jasné formulace požadavků pro jednotlivé dovednosti, v některých oblastech příliš vysoké požadavky, ověřování znalosti reálií v rámci jazykové zkoušky bez adekvátního vysvětlení způsobu jejich zjišťování a podílu na celkovém hodnocení a v neposlední řadě důraz na formulaci požadavků na receptivní dovednosti - požadavky na výkony v produktivních dovednostech jsou formulovány velmi obecně a typy možných testových úloh nejsou vůbec uvedeny.

Je rovněž prezentován autorčin návrh alternativního katalogu požadavků k maturitní zkoušce z anglického jazyka pro úroveň mírně a středně pokročilou. Návrh požadavků vychází z popisů jazykových úrovní Společného evropského referenčního rámce pro jazyky. Alternativní katalog obsahuje formulace požadavků na základní řečové dovednosti (poslech s porozuměním, čtení s porozuměním, psaní a mluvení), jazykové znalosti (slovní zásoba, struktura jazyka) pro obě uvedené úrovně obtížnosti, jakož i hodnotící kritéria pro produktivní řečové dovednosti. U požadavků na receptivní dovednosti je stanoven požadovaný výkon – co kandidát na dané úrovni a v dané řečové dovednosti umí (např. rozeznat, identifikovat, odvodit apod.) – a podmínky, ve kterých kandidát tento výkon provede (např. kandidát nepoužívá slovník, test probíhá v jazykové laboratoři apod.). Pro jednotlivé produktivní dovednosti je podobně jako pro receptivní dovednosti formulován výkon (např. formulovat myšlenky, názor a otázky v delších a složitějších souvětích a v abstraktní rovině; podat jasné a detailní popisy apod.). Na rozdíl od receptivních dovedností však nejsou specifikovány podmínky výkonu, ale je stanoveno, jak dobře musí kandidát výkon provést, aby odpovídal požadovaným kritériím (např. kandidát mluví v krátkých, jednoduchých větách; může se dopouštět chyb a nepřesností i v používání jednoduchých jazykových prostředků apod.). Uvádějí se požadované jazykové prostředky, slovní zásoba a funkce pro obě úrovně obtížnosti. Jsou uvedeny doporučené testovací techniky a hodnotící kritéria produktivních dovedností.

Tento návrh alternativního katalogu byl posouzen několika rodiči a učiteli, jejichž připomínky jsou v této části práce rovněž uvedeny a komentovány. Dle vyplněných dotazníků se všem respondentům (9) jeví formulace požadavků srozumitelné. Neshodují se v názorech na časové limity, ale tento problém je možné kvalifikovaně posoudit pouze po pretestování. Středoškolští pedagogové se domnívají, že je sice důležité naučit žáky cizojazyčné komunikaci v každodenním životě, nicméně poukazují na skutečnost, že musejí mnoho času věnovat i gramatice, protože u přijímacích zkoušek na vysoké školy je toto více-méně jediný způsob ověřování jazykové úrovně uchazečů. Tuto skutečnost lze hodnotit jako negativní.

Nedílnou součástí druhé části práce je i pasáž týkající se procesu konstrukce dvou dílčích testů ověřujících úroveň čtení a psaní pro úroveň mírně pokročilou. Specifikační tabulky pro jednotlivé subtesty shrnují jejich nejdůležitější charakteristiky. Tyto subtesty prošly běžným připomínkovým procesem na pracovišti autorky, na němž se podíleli dva pracovníci teoreticky připravení a zkušení v praxi jazykového testování. Do připomínkového řízení nebyli zapojeni středoškolští učitelé angličtiny, protože přestože jsou pro svou práci kvalifikováni, většinou nejsou proškoleni pro tvorbu testů.

Po zvážení a zařazení relevantních připomínek byly oba dílčí testy postupně zadávány na pěti středních školách (čtyři gymnáziá a jedna střední odborná škola) a na jedné vysoké škole. Celkově byly oba nebo pouze jeden subtest zadány 126 kandidátům: oba subtesty napsalo 85 kandidátů, subtest čtení 110 a subtest psaní 101 kandidátů. Testy byly zadávány v období listopad 2005 – únor 2006.

Subtest čtení byl hodnocen výhradně autorkou, protože učitelé středních škol hodnocení odmítli z důvodu vysokého pracovního vytížení. Pretestování odhalilo jisté formální

nedostatky testu (nedostačující instrukce, nedostatek místa pro psaní odpovědi atd.). Někteří středoškolští učitelé také kritizovali výskyt neznámých slov v textu – tato výhrada nebyla přijata, protože jedním z cílů testu bylo ověřit, jestli by kandidáti v reálném životě porozuměli autentickému textu, který pro jejich potřeby nikdo neupravuje. Navíc jednotlivé testové položky byly záměrně jednodušší, aby obtížnost textu samotného vyvážily. Protože hodnocení jednotlivých položek bylo většinou objektivní (18 z 30 položek hodnoceno objektivně, 9 doplňovacích a 3 otázky s krátkými odpověďmi), při opravě testu nedošlo k závažným problémům. Připomínky zadávajících učitelů i autorky byly sumarizovány a bylo k nim přihlíženo.

Subtest psaní byl hodnocen dvěma nezávislými hodnotiteli – jedním z nich byla vždy autorka a druhým kolega z pracoviště autorky – tato skutečnost opět vyplynula z neochoty středoškolských pedagogů na hodnocení participovat. Hodnocení bylo provedeno vždy podle kritérií hodnocení jednotlivých testových úloh a nezávisle. Hodnotitelé spolu konzultovali pouze hodnocení, ve kterých se jednalo o počty bodů lišící se vzájemně o více než dva body; v těchto případech hodnotitelé prošli ještě jednou jak vypracované testové úlohy, tak kritéria hodnocení a dospěli ke kompromisnímu řešení. Průměr rozdílů v hodnocení všech subtestů psaní byl 1,1 bodu a korelace hodnocení autorky s ostatními hodnotiteli je 0,96. Tento výsledek lze pokládat za velmi úspěšný.

Výsledky subtestu čtení a psaní od kandidátů, kteří psali oba subtesty byly vzájemně korelovány. Hodnota korelace dosahuje 0,47, což naznačuje, že každý subtest ověřuje jiné dovednosti a tedy používání obou subtestů je opodstatněné.

Pro důkladnou položkovou analýzu dílčího testu čtení s porozuměním byl použit počítačový software ITEMAN, který pro každou položku vyhodnotil její obtížnost, diskriminační index a v případě položek s výběrem z vícenásobné volby také fungování distraktorů. Ve většině případů splňovaly tyto parametry běžné požadavky. Tento subtest obsahuje 30 položek, které byly hodnoceny dichotomicky. Průměrné skóre (*mean*) u testu bylo 20,73. Standardní odchylka se rovná 5,40. Reliabilita (*reliability*) dosáhla hodnoty 0,844. Rozložení u analyzovaného testu je bimodální, což je pro potřeby rozřazení kandidátů na úspěšné a neúspěšné vhodné.

Práce vychází ze systematického a koncepčně propracovaného přístupu k problematice požadavků k maturitní zkoušce z anglického jazyka. I když nemohla postihnout všechny aspekty dané problematiky, nabízí ucelený model podpořený empiricky ověřenými daty s dostatečným prostorem pro případné modifikace.

### **Směry dalšího výzkumu**

Výsledkem disertační práce je návrh alternativního katalogu požadavků zjišťujícího všechny řečové dovednosti. Žádný výzkum menšího rozsahu však nemůže poskytnout zcela vyčerpávající výsledky a proto k dalšímu výzkumu lze doporučit následující.

Bylo by vhodné nechat alternativní katalog posoudit větším počtem učitelů středních i vysokých škol, čímž by bylo možné zprostředkovat širší zpětnou vazbu. Na základě vyhodnocení výsledků kvantitativního výzkumu by bylo užitečné provést strukturované nebo zčásti strukturované rozhovory k problematickým otázkám a tím doplnit zpětnou vazbu kvalitativním výzkumem.

Nejvíce pozornosti budou vyžadovat testy, a to nejen subtesty poslechu s porozuměním a mluvením, ale i dva již na menším počtu studentů ověřované subtesty čtení a psaní.

Při zadání subtestu poslechu by nahrávka měla odpovídat tempu a srozumitelnosti, které uvádí alternativní katalog. Výsledky subtestu by měly být také podrobeny položkové analýze. Pilotování testu mluvení bude pak vyžadovat systematické pozorování a vyhodnocení vhodnosti testových úloh. V praxi bude nezbytné zaměřit se jak na způsob vedení ústní zkoušky, tak na uplatnění hodnotících kritérií. Pro hodnocení psaní i mluvení bude potřeba hodnotitele vyškolit.



## PŘÍLOHY

### Příloha P – 1 Pojmy používané v disertační práci a jejich definice

#### **dílčí dovednost (*sub-skill, micro skill*)**

„základní řečové dovednosti se často rozdělují na dílčí dovednosti, jako rozlišování zvuků v poslechovém testu, signalizace vztahů mezi různými částmi výpovědi v mluveném nebo psaném projevu, identifikování relevantní informace v textu. Testové položky mohou být zaměřeny na jednu nebo více dílčích dovedností“ (Davies et al., 1999, s. 180).

#### **distraktor (*distractor*)**

„jakákoliv odpověď u uzavřených testových úloh, která není správná, to znamená, že to není správná odpověď, ale je nabízená jako prostředek zjištění, jestli jsou kandidáti schopni odlišit správnou odpověď od několika alternativ“ (Davies et al., 1999, s. 49).

#### **dílčí test (*sub-test*)**

viz subtest

#### **gramaticko-překladová metoda (*grammar-translation method*)**

metoda používaná ve výuce jazyků zejména v minulosti; zaměřovala se spíše na vyučování „o jazyce“ než na vyučování „jazyku“; tato metoda dává důraz na učení gramatickým pravidlům a nejčastěji používanými cvičeními jsou překlady z a do cílového jazyka.

#### **hodnocení (*assessment*)**

„jakákoli metoda systematického získávání informací z testů a jiných zdrojů, vedoucí k závěrům o charakteristikách lidí, věcí nebo programů“ (Standards pro pedagogické a psychologické testování, 2001, s.174)

#### **hodnotící kritéria, kritéria hodnocení (*assessment criteria*)**

„kvality (vlastnosti), podle kterých se posuzuje výkon u testu. Úspěšný výkon u hodnocené úlohy může být charakterizován pomocí lingvistických i nelingvistických kritérií“ (Davies et al., 1999, s. 38).

#### **hodnotící škála (*rating scale, proficiency scale*)**

„škála určena k popisu dosažených jazykových znalostí a řečových dovedností obsahující sérii definovaných úrovní, vůči kterým se posuzuje výkon učícího se.“ (Davies et al., 1999, s.153).

#### **hodnotitel (*examiner, rater, assessor*)**

osoba pověřená hodnocením výkonů u zkoušky a speciálně pro tento účel vyškolená.

#### **inference (*inference*)**

„předpoklad vytvořený o schopnosti kandidáta na základě jeho/jejího výkonu u testu“ (Davies et al., 1999, s. 81).

**jazykové prostředky (*language tools*)**

nástroje potřebné pro realizování určitého komunikačního záměru; tyto nástroje mohou být na úrovni lexikální, gramatické, syntaktické, pragmatické, funkční, sociolingvistické apod.

**jazykové znalosti (*language knowledge*)**

znalosti jazykových prostředků

**kandidát (*candidate, cand*)**

osoba podstupující test nebo zkoušku

**koeficient alfa (*coefficient alpha*)**

„koeficient vnitřní konzistence testu, založený na počtu částí, ze kterých se test skládá (např. položky, subtesty nebo posuzovatelé), na vzájemných vztazích těchto částí a na rozptylu celkového skóru. Bývá označován také jako Cronbachovo alfa ...“ (Standardy pro pedagogické a psychologické testování, 2001, s.175).

**komunikativní přístupy (*communicative approaches*)**

přístupy k výuce a testování cizího jazyka, jejichž cílem je rozvíjet a hodnotit schopnost v daném jazyce komunikovat.

**konstrukt (*construct*)**

„teoretický koncept nebo vlastnost, kterou má test měřit“ (Standardy pro pedagogické a psychologické testování, 2001, s.176).

**objektivní bodování (*objective scoring/marking*)**

„bodovací postupy určeny pro testové položky, které nevyžadují od hodnotitelů subjektivní rozhodování. Všechny přijatelné odpovědi jsou jasně specifikovány v bodovacím klíči“ (Davies et al., 1999, s.132).

**pretestování, také pilotáž (*pre-testing, piloting*)**

„administrace testu vzorku testovaných osob s cílem vyzkoušet určité aspekty testu nebo testových položek, jako jsou instrukce, časové limity, formát a možnosti odpovědí na položky“ (Standardy pro pedagogické a psychologické testování, 2001, s.179).

**produktivní řečové dovednosti (*productive skills*)**

„řečové dovednosti mluvení a psaní. Produktivní řečové dovednosti jsou přímo pozorovatelné v testovém výkonu, na rozdíl od receptivních řečových dovedností čtení a poslechu, a obvykle se hodnotí subjektivně podle těch kritérií, která jsou pro účel testu pokládána za důležitá“ (Davies et al., 1999, s. 152).

**receptivní řečové dovednosti (*receptive skills*)**

„řečové dovednosti poslechu a čtení. Na rozdíl od produktivních řečových dovedností psaní a mluvení je nelze přímo pozorovat v testovém výkonu, protože procesy pochopení (porozumění) probíhají v mysli. Z tohoto důvodu jsou kandidáti často požádáni

vyprodukovat určitý typ psaného nebo mluveného projevu, aby prokázali pochopení (porozumění)“ (Davies et al., 1999, s.164).

#### **reliabilita (*reliability*)**

„reliabilita je míra, do níž jsou výsledky testu konzistentní; kdyby kandidáti dělali stejný test dnes i zítra, dosáhli by stejných výsledků (za předpokladu, že nedošlo ke změnám v jejich schopnostech)? Existuje několik způsobů měření reliability objektivních testů (opakování testu/*test-retest*; paralelní testy/*parallel forms*; půlení/*split-half*; KR 20, KR 21 atd.). Reliabilita subjektivních testů se měří výpočtem spolehlivosti bodování.“ (Anderson; Clapham; Wall 1995, s. 294).

„míra, ve které jsou testové skóry ve skupině testovaných osob konzistentní při opakované aplikaci postupů měření a proto spolehlivé a opakovatelné u jednotlivých testovaných osob“ (Standardy pro pedagogické a psychologické testování, 2001, s.181).

#### **řečová dovednost (*language skill, skill, macro-skill*)**

„jeden z aspektů, které jsou základem pro používání jazyka. Jazykové testy se často charakterizují podle toho, která/é ze čtyř dovedností (také známých jako makro dovednosti) jsou obsaženy ve výkonu; to znamená, že se běžně mluví o testech čtení, psaní, poslechu a mluvení“ (Davies et al., 1999, s.180); schopnost používat jazykové prostředky v praktické komunikaci.

#### **sehrání situační role (*role play*)**

„úloha v testu ústního projevu, ve které kandidát přijímá určitou roli v interakci s dalším jedním nebo více mluvčími (například s dalším kandidátem, učitelem nebo vyškoleným zkoušejícím). Role a úloha jsou stanoveny v instrukcích, které účastníci dostanou buď před testem nebo na začátku testu. Obecně tyto role simulují autentické situace relevantní pro komunikační požadavky, ve kterých se kandidát může nacházet v reálném životě“ (Davies et al., 1999, s. 171).

#### **spravedlivý přístup (*fairness*)**

„při testování je to takový princip, že každé testované osobě by se mělo dostat stejného zacházení“ (Standardy pro pedagogické a psychologické testování, 2001, s.182).

#### **standardizovaný test (*standardized test*)**

test, který má za ideálních podmínek následující charakteristiky, z nichž ne všechny jsou vždy dodrženy v jazykových testech, které se označují jako standardizované:

- přísně dodržovaný proces vývoje, ověřování a revidování, který určuje měřicí vlastnosti testu;
- standardní postupy pro zadávání a bodování testů;
- standardní obsah testu ve všech verzích. Tento obsah je založen na souboru testových specifikací (kráceno: Davies et al., 1999, s. 187).

#### **subjektivní bodování (*subjective scoring/marking*)**

„subjektivní bodování se obvykle používá pro bodování testů psaní nebo mluvení. Od hodnotitelů se požadují rozhodnutí, která jsou složitější než rozhodnutí správné –

nesprávně... mají za úkol zhodnotit, jak dobře kandidát splnil daný úkol, k čemuž potřebují hodnotící škálu“ (Anderson; Clapham; Wall 1995. s. 107).

#### **subtest (*sub-test*)**

dílčí test;

„samostatná část baterie testů, která je věnována hodnocení určité dovednosti, jako například gramatiky, mluvení nebo čtení. Počty bodů z jednotlivých subtestů se někdy uvádějí samostatně, někdy spolu s celkovým počtem bodů, který vyjadřuje celkový výkon v testu, nebo jsou uváděny jako jeho alternativa“ (Davies et al., 1999, s. 191).

#### **technika testování (*testing technique*)**

způsob testování; technika používaná na zjišťování úrovně jazykových znalostí a/nebo dovedností. Testovací techniky jsou například doplňování chybějících slov/výrazů do textu (*gap filling*), volba z nabídky odpovědí, nejčastěji ze čtyř (*multiple choice*), přiřazování (*matching*), popis obrázku (*picture description*) apod.

#### **test (*test*)**

jakýkoliv postup, jehož cílem je měření schopnosti, znalostí nebo výkonu (Richards; Platt; Platt 1992, s. 377), viz zkouška;

„systematický postup určený k pozorování chování jednotlivce a k jeho popisu s pomocí číselné škály nebo kategorizačního systému“ (Cronbach, in McNamara, 1996, s.10).

#### **test dosažené úrovně (*proficiency test*), test způsobilosti**

„test, který měří, kolik se někdo z jazyka naučil. Na rozdíl od testů ověřujících zvládnutí učiva, test dosažené úrovně není založen na konkrétním výukovém kurzu. Test dosažené úrovně často měří, co se kandidát naučil ve vztahu k určitému požadavku reálného světa, např. zvládá cílový jazyk dostatečně na to, aby sledoval/a přednášku, učil/a se inženýrem/inženýrkou, pracoval/a jako lyžařský instruktor/ka v tomto jazyce, nebo překládal/a z tohoto jazyka. Některé testy dosažené úrovně byly standardizovány za účelem celosvětového použití, například americký test TOEFL, který se používá k měření dosažené úrovně anglického jazyka u zahraničních studentů vyššího vzdělávání, kteří chtějí studovat v USA; nebo britsko-australský test IELTS navržený pro ty, kdo chtějí studovat ve Velké Británii nebo v Austrálii“ (Davies et al., 1999, s. 154).

#### **test ověřující zvládnutí učiva (*achievement test, attainment test*)**

„vztahuje se ke zvládnutí toho, co se kandidát naučil, co bylo vyučováno a co je obsaženo v osnovách, učebnicích, výukových materiálech“ (Davies et al., 1999, s. 2).

#### **testovací technika (*testing technique*)**

viz technika testování

#### **testová položka (*test item*)**

„ta část testu, která požaduje od kandidáta specifikovanou formu odpovědi. Na základě této odpovědi nebo souboru odpovědí se rozhoduje o kandidátově znalosti nebo schopnosti. Testové položky mohou být nejrůznějších druhů, a to podle typu a účelu testu“ (Davies et al., 1999, s. 201);

„význam termínů testová úloha (*task*) a testová položka (*test item*) se překrývá. Pokud se mezi nimi rozdíl dělá, obvykle to implikuje, že položka je menší a méně komplikovaná, zatímco úloha je rozsáhlejší a složitější. Test obvykle obsahuje buď větší počet relativně krátkých položek nebo malý počet relativně složitých úloh“ (Davies et al., 1999, s. 196).

#### **testová úloha (*task*)**

„to, co se od kandidáta požaduje, aby udělal během testu nebo jeho části, jako například psaní poznámek z mluveného textu, účast na interview nebo sehrání situační role, čtení textu a zodpovězení otázek týkajících se porozumění textu“ (Davies et al., 1999, s. 196).

#### **testová úloha, otevřená (*open-ended task, constructed-response item, free response item*)**

„typ testové položky nebo úlohy, který vyžaduje od kandidátů, aby formulovali své vlastní odpovědi a nevybírali je z poskytnuté nabídky odpovědí. Jedním z účelů použití otevřených testových úloh je umožnit hodnocení schopnosti kandidátů produkovat jazyk – proto přímé testy mluvení nebo psaní vždy vyžadují otevřené testové úlohy“ (Davies et al., 1999, s. 32).

#### **testová úloha, uzavřená (*forced-choice item, fixed-response item, closed-ended response*)**

„testová úloha, ve které musí být správná odpověď vybrána z několika alternativ“ (Richards; Platt; Platt 1999, s. 377).

#### **testové specifikace (*test specifications, specs*)**

„dokument, který stanovuje, co má test měřit a jak to bude testovat. Zatímco základním adresátem testových specifikací jsou tvůrci testů/položek, mohou být specifikace (nebo jejich modifikovaná forma) také určeny hodnotitelům testů a uživatelům testů“ (Davies et al., 1999, s. 207).

#### **úroveň (*level*)**

„1. míra, stupeň (např. 1 – 9, nebo A – E) nebo popis způsobilosti nebo schopnosti kandidáta, obvykle popsána na určitém druhu škály a určena na základě výkonu u testu.“  
„2. označení nebo popis obtížnosti testu ...(například začátečník, mírně pokročilý, pokročilý) nebo úloh či textů, které test obsahuje...“ (Davies et al., 1999, s. 107).

#### **validita (*validity*)**

„validita se obecně vztahuje k vhodnosti určitého testu nebo jakékoliv jeho součásti jako měřítka toho, co se má měřit. O testu se říká, že je validní do té míry, do které měří to, co se od něho očekává, aby měřil. Z toho plyne, že termín *validní*, pokud se používá k popisu testu, by měl být doprovázen předložkou *pro*. Jakýkoliv test pak může být validní pro určité účely, ale ne pro účely jiné“ (Henning, 1987, in Alderson; Clapham; Wall 1995, s. 170).

„validita je definovaná jako míra, ve které empirické důkazy a teorie podporují interpretaci testových skóre při doporučeném použití testu. Validita je proto nejdůležitějším hlediskem při tvorbě a hodnocení testu“ (Standards pro pedagogické a psychologické testování, 2001, s. 17).

**výkon (*performance*)**

„chování, které projevuje kandidát během plnění určitého úkolu. ... Zatímco hodnocení schopnosti je založeno na tomto pozorovatelném chování, bere se v úvahu, že aspekty testové situace mohou způsobit, že kandidát projevuje svůj výkon takovým způsobem, který nedovoluje získat přesné měřítko jeho schopnosti“ (Davies et al., 1999, s. 143).

**zkouška (*examination, exam*)**

jákykoliv postup, jehož cílem je měření schopnosti, znalostí nebo výkonu (Richards; Platt; Platt 1992, s. 377), viz test

**zpětný efekt/vliv (*backwash* nebo *washback effect*) určité zkoušky na vyučování**

„vliv testování na výuku. Zpětný efekt jazykových testů může být pozitivní nebo negativní. O existenci negativního zpětného vlivu se hovoří, když jsou testové položky založeny na zastaralém chápání jazyka, které jenom málo souvisí s učebními osnovami. Když se například dovednost psaní testuje pouze položkami volby z nabídky odpovědí, potom existuje velký tlak procvičovat spíše takovéto položky než procvičovat samotnou dovednost psaní. ...O pozitivním zpětném vlivu se hovoří v tom případě, když se zavádějí takové testovací postupy, které podporují „dobré“ metody vyučování. Například použití ústního rozhovoru u závěrečné zkoušky může vést učitele k tomu, aby se žáky procvičovali používání konverzačního jazyka. ...Bylo poukázáno na skutečnost, že pro mnohá tvrzení týkající se pozitivního a negativního dopadu jazykového testování na vyučování existuje pouze málo empirických důkazů“ (Davies et al., 1999, s. 225).

## Příloha P – 2 Výsledky pozorování průběhu maturitní zkoušky

### Forma maturitní zkoušky

Na žádné z pozorovaných škol neexistuje oficiálně písemná část maturitní zkoušky. Tudíž tato zkouška vysoké důležitosti nezjišťuje systematicky a objektivně úroveň dovednosti poslechu a čtení s porozuměním. Stejně tak se nezjišťuje ani úroveň produktivní dovednosti psaní. Je nakonec však také sporné, jestli zjišťuje úroveň dovednosti mluvení, nebo pouze prezentaci namemorovaných cizojazyčných konverzačních témat a reálií.

Ve formě maturity se objevují některé charakteristiky společné pro všechny pozorované školy. Zkoušený si vytáhne otázku z tematických okruhů k maturitě. U většiny škol k tematickému okruhu také dostane text, který je stejného obsahového zaměření. U některých maturit se tento text čte nahlas a pak se odpoví na otázky za textem, které zjišťují porozumění tohoto textu. U jiných škol se text nahlas nečte, pouze se odpovídá na otázky k textu. Kromě dvou škol, kde vyučující přiznávají, že se od maturanta očekává spíše monologický projev, se všichni ostatní vyučující a tedy i zkoušející shodují v názoru, že u maturity očekávají spíše dialog než monolog. Na dvou školách se také zkouší gramatika – na jedné škole pouze okrajově a vychází se z textu, na druhé se zadá samostatná gramatická otázka a úkolem maturanta je vysvětlit určitý gramatický jev a demonstrovat jeho aplikaci na doplňovacím cvičení.

Na žádné ze středních škol neexistují kritéria pro hodnocení ústního projevu. Podle slov dotazovaných učitelů se jedná o hodnocení celkové, často se přihlíží k práci během studia. Není však přesně určené, kolik procent známky určuje například úroveň přípravy do hodin. Také není ani učitelům samotným zcela jasné, jestli hodnotí úroveň jazyka nebo faktografických znalostí. Kromě toho neexistuje jednotná koncepce maturitní zkoušky ani v rámci jedné školy. Je v pravomoci každého vyučujícího připravit si maturitní otázky podle svého názoru. Takový přístup k dané problematice dává jasně tušit, že výsledky u maturity z anglického jazyka budou zřejmě těžko srovnatelné dokonce i v rámci jedné školy, natož u různých škol. Je zjevné, že učitelé ani nejsou vedeni k uvažování nad tím, co vlastně chtějí u maturitní zkoušky zjistit a tudíž také neuvažují nad tím, jakého nástroje k tomuto zjištění použít.

Na všech školách bylo vypracováno 25 maturitních otázek. Tyto otázky ve všech školách byly silnou motivací i určitým vodítkem jak pro učitele, tak pro žáky. Na některých školách se připravovali žáci na otázky samostatně a v hodinách se učitelé těmito tématy zabývali, na dvou školách učitelé více-méně vypracovali otázky a jediné, co se po žácích chtělo bylo, aby se tyto otázky tzv. naučili.

Na jedné odborné škole měly otázky 2 verze – v jedné z nich se vyskytly tři otázky s odbornějším zaměřením (*PC hardware, PC software, PC applications*). Tato verze byla připravena pro třídu se zaměřením na počítače. Na obou dalších odborných školách se v prvním případě v jedné a v druhém případě ve dvou otázkách objevilo odborné zaměření škol (*SPSCH – Chemistry – a science and profession, SPŠTX – Garments and*

*clothes, Some facts about fashion*). Otázky ze čtyř škol byly formulovány velmi stručně (*Housing, Canada, Health and Diseases...*). Naopak otázky ze dvou dalších škol obsahovaly také určité stimuly, které mohou žákovi pomoci při formulaci myšlenek (*My biography, Hobbies, Leisure time activities and plans for the future: your daily programme on weekdays and at weekends, talk about your family, favourite subjects at school, ...*).

Všechny školy kromě jedné používají monotematickou orientaci otázek, to znamená, že žák si vytáhne téma, o kterém má mluvit, a také k němu dostane text vztahující se ke stejné problematice. Proti tomuto přístupu je možné namítnout, že vzhledem k tomu, že test může měřit pouze jistý vzorek probraného učiva nebo látky explicitně vyjádřené v testových specifikacích, by tento vzorek měl co možno nejlépe reprezentovat to, co chceme hodnotit. Monotematické zaměření obou otázek je teda v rozporu s tímto požadavkem.

Pouze jedna škola připravila otázky, které nepokrývají pouze jeden tematický okruh. V těchto otázkách je první část z oblasti konverzačních okruhů, zatímco druhá část je zaměřena na otázky pokrývající realie, dějiny a literaturu anglicky mluvících zemí. To znamená, že počet otázek fakticky nebyl 25 jako u jiných škol, nýbrž 50. V tomto případě zkoušení žáci neobdrželi žádný text, nýbrž celá zkouška byla zčásti monologem, zčásti dialogem. Můžeme říct, že tato škola byla daleko více „přátelská“ ke zkoušeným a zároveň se nevědomě snažila o získání reprezentativnějšího vzorku projevu žáků než ostatní školy.

Jeden z přístupů k tvorbě jazykových zkoušek a jejich používání vychází z těchto dvou základních principů:

- Je nutné, aby výkon u zkoušky (*performance*) prokazatelně korespondoval s používáním jazyka v mimo testových situacích.
- Pro kontrolu kvality testování je důležité jasně a explicitně definovat vlastnosti užitečnosti/výhodnosti testu (*test usefulness*), jako je reliabilita, konstruktová validita, autenticita, interaktivnost, dopad a praktičnost. (L.F.Bachman; A.S.Palmer 1996, s. 300 - 301).

Stačí se pouze zběžně seznámit s tím, jak v praxi probíhají maturitní zkoušky z angličtiny, a je zcela evidentní, že ani jeden z výše uvedených principů nebyl brán v úvahu. Dovedeme si snad představit situaci v reálném životě, kdy by žák měl v monologickém projevu v cizím jazyce vyjmenovat čtyři roční období, různé nemoci, rozličné druhy obchodů, oděvů, sportů, jídel? Je určité vhodné, pokud zmiňovanou slovní zásobu zná, ale tuto znalost, která má sloužit jako prostředek při komunikaci, nelze zaměňovat za znalost cílovou. Způsob zkoušení u maturitní zkoušky bohužel jasně odráží širší problém, který se už několik let odváženě kritizuje, nicméně stále přetrvává – je to důraz na teoretické, encyklopedické a faktografické znalosti místo akcentování schopnosti praktického použití znalostí. Bílá kniha také na tento problém v obecné rovině poukazuje – kromě jiného například doporučuje „podporovat rozvoj klíčových kompetencí jako nástroje přeměny encyklopedického pojetí vzdělávání“.



## Pozorování maturity

Na každé škole byla ústní maturitní zkouška obsahem i formou odlišná. Společným pro všechny školy naopak bylo, že žáci v určité etapě zkoušky (někde v první, někde ve druhé části maturitní zkoušky) prezentovali *monologický projev* na určité předem vybrané téma. Někteří žáci zjevně prezentovali z paměti naučené fakta (školský systém ve Velké Británii, různé svátky v ČR a Velké Británii, Kanada, nákupy, zdraví a nemoci, životopis, Londýn...), u jiných naopak bylo zjevné, že i když znali otázky předem a určitě se na ně nějakým způsobem připravovali, přece působili přirozeně a budili dojem, že mluví i o vlastních názorech a skutečnostech. Z uvedeného je možné také usuzovat, jak v hodinách anglického jazyka probíhá výuka a k čemu jsou žáci vedeni. Také to ukazuje, že i když zadání úlohy je velmi podobné, může vést k velmi odlišné produkci ústního projevu. Nutno podotknout, že na různých školách zkoušející různě do tohoto monologického projevu zasahovali. Většinou šlo o velmi vhodné vstupování, čímž mezi zkoušenými a zkoušejícími vznikla interakce. Proto lze spíše mluvit o dialogu. Ne všechny otázky však byly položeny zcela vhodně – blíže o tom pojednává podnadpis *Typy otázek*.

Pouze na jedné střední škole nebylo kromě výše uvedeného způsobu zkoušení použito například *textu ke čtení*. Tento text byl naopak použitý v ostatních školách, a to zejména k zjištění porozumění čtenému. Na jedné škole se zkoušející ptala *otázky k textu*. Na dvou školách šlo o *doslovný překlad* textu, a to překlad po větách. Na dvou školách bylo součástí maturitní zkoušky *také zkoušení gramatiky*. Proč se kombinuje hodnocení ústního projevu, a u některých otázek také znalost reálií nebo jiných fakt, s hodnocením čtení s porozuměním, zůstává nejasné.

### *Typy otázek*

Je důležité podotknout, že typy otázek, které kladli zkoušející, byly téměř výhradně vysoce pravděpodobné, takže zkoušení je mohli očekávat. Otázky se často zaměřovaly na zkoušení slovní zásoby, faktografických znalostí a ve dvou případech i na znalosti gramatiky. Otázky jiného druhu úzce souvisely s daným tématem, často to byly tzv. *Yes – no questions*, tedy otázky, na které je možné odpovědět ano nebo ne. Tento typ otázek by se měl u ústních jazykových zkoušek minimalizovat, protože nestimuluje zkoušeného k produkci hodnotitelného vzorku svého ústního projevu. Nelze říct, že tyto otázky se vůbec nemají používat. Nicméně důležité je, aby za takovými otázkami následovaly otázky navazující. Pozitivní stránkou u všech maturit bylo, že byly vedeny v cílovém, tj. anglickém jazyce, až na dvě školy, u kterých zkouška obsahovala také část zameranou na zkoušení gramatiky – tady měli žáci objasnit daný gramatický jev v češtině.

Uvedené příklady otázek byly v anglickém jazyce.

Pro ilustraci bych ráda uvedla několik příkladů otázek, které:

A/ zkoušely slovní zásobu:

*Vyjmenujte různé typy bydlení ve Velké Británii, jaký obchod prodává zboží, které vidíte na obrázku, kam byste šel, kdybyste potřeboval koupit oblek, počítačovou hru..., kam jdete, když potřebujete spravit boty, auto..., co mají lidé na obrázku na sobě, co obvykle máte ke snídani, obědu, večeři, kde můžete koupit dort, jak se jmenují menší specializované obchody, co můžeme koupit v drogerii, jak řekneme anglicky leštidla, jaké jsou hlavní části počítače, jaké druhy tiskáren znáte.*

B/ zkoušely faktografické znalosti:

*Kdy začíná v Británii školní docházka, jaké univerzity v Británii znáte, jaké je hlavní město Kanady, jaký je počet obyvatelů Kanady, vyjmenujte hlavní jezera v Kanadě, z čeho pochází jméno Kanada, jaký je tam politický systém, popište kanadskou vlajku, čím byl původně Tower, čím je dnes, jak se jmenuje současná britská královna, jaká řeka protéká Londýnem, jaké známé mosty přes Temži znáte, čím je známý Tower Bridge, jaké jsou hlavní body životopisu, co jí Britové k obědu.*

C/ byly otázkami, na které se dalo odpovědět ano-ne a nenásledovaly za nimi otázky další:

*Chodíte na koncerty, chtěla byste pracovat jako průvodkyně, pracovala byste ráda s lidmi, potřebujete angličtinu pro práci s počítači, jste spokojen se školou, jsou učebny dobře vybaveny, máte ráda sport, díváte se ráda na sport, zúčastňovala jste se sportovních událostí na škole, chtěla byste navštívit USA.*

D/ byly vhodně položené:

*Jaký je rozdíl mezi bydlením v ČR a v Británii, viděla jste nějaký muzikál také v Praze, srovnajte toto představení s muzikálem, který jste viděla v Londýně, chtěl byste bydlet v domku samostatném nebo řadovém, proč, bydlel byste raději ve městě nebo na vesnici a proč, jak si představujete ideální bydlení.*

## **Závěr**

Z uvedeného průzkumu vyplývá, že na pozorovaných středních školách se forma i obsah maturitních zkoušek značně liší, z čehož je zřejmé, že se bude lišit i způsob hodnocení úrovně maturantů. Tato situace je důsledkem toho, že nejsou explicitně specifikovány požadavky, které musí být u zkoušky splněny. Kandidáti jsou zkoušeni svými vyučujícími a do hodnocení podvědomě zahrnují i práci studenta během studia. Nejsou ani rámcově formulována kritéria hodnocení ústního projevu. Učitelé si vlastně ani neuvědomují, které dovednosti zkoušejí a jaká kritéria hodnocení používají. Je vysoce pravděpodobné, že stejný kandidát by nebyl stejně hodnocen na různých školách. Za tuto situaci nelze vinit učitele, protože připravují a vedou maturitní zkoušky tak, jak si je pamatují ze svých studií a tak, jak je to naučili jejich kolegové. Navíc k testování jim žádné školení poskytnuto nebylo. Velmi podobnou situaci zaznamenal při svém detailním pozorování a analýze Charles Alderson, který působil jako specialista na jazykové testování v Maďarsku (Alderson; Nagy; Oveges 2000).

## Příloha P – 3 Dotazník k hodnocení alternativního katalogu požadavků

### Hodnocení alternativního návrhu Katalogů požadavků pro maturitní zkoušku z anglického jazyka

Přečtěte si prosím alternativní návrh Katalogu požadavků pro maturitní zkoušku z anglického jazyka. Poté odpovězte na otázky 1-8. Výběr z nabízených možností proveďte zakroužkováním a pokud k tomu budete vyzváni, svoji volbu zdůvodněte. Vaše odpovědi jsou anonymní.

#### 1. Je účel maturitní zkoušky popsán srozumitelně?

ANO - NE

Pokud odpovíte ne, specifikujte, které formulace Vám nejsou jasné.

---

---

---

#### 2. Je obecná charakteristika obou požadovaných jazykových úrovní srozumitelná?

ANO - NE

Pokud odpovíte ne, specifikujte, které formulace Vám nejsou jasné.

---

---

---

#### 3. Dokážete si po přečtení požadavků na jednotlivé řečové dovednosti představit, co se od žáků u zkoušky bude požadovat?

ANO - NE

Pokud nedokážete, specifikujte, které formulace Vám nejsou jasné.

---

---

---

#### 4. Jak byste v tomto alternativním návrhu Katalogu požadavků celkově zhodnotil/a požadavky pro zkoušky na jazykové úrovni „mírně pokročilý“?

a/ Mají odpovídající náročnost.

b/ Jsou příliš málo náročné.

Uveďte proč.

---

---

---

c/ Jsou příliš náročné.

Uveďte proč.

---

---

---

5. Jak byste v tomto alternativním návrhu Katalogu požadavků celkově zhodnotil/a požadavky pro zkoušky na jazykové úrovni „středně pokročilý“?

a/ Mají odpovídající náročnost.

b/ Jsou příliš **málo** náročné.  
Uveďte proč.

---

---

---

c/ Jsou **příliš** náročné.  
Uveďte proč.

---

---

---

6. U zkoušky pro úroveň „mírně pokročilý“, souhlasíte s

a/ počtem dílčích testů? ANO - NE  
Pokud nesouhlasíte, uveďte proč.

---

---

---

b/ časovými limity pro jednotlivé dílčí testy? ANO - NE  
Pokud nesouhlasíte, uveďte proč.

---

---

---

c/ počty bodů pro jednotlivé dílčí testy? ANO - NE  
Pokud nesouhlasíte, uveďte proč.

---

---

---

7. U zkoušky pro úroveň „středně pokročilý“, souhlasíte s

a/ počtem dílčích testů? ANO - NE  
Pokud nesouhlasíte, uveďte proč

---

---

---

b/ časovými limity pro jednotlivé dílčí testy? ANO - NE  
Pokud nesouhlasíte, uveďte proč

---

---

---

c/ počty bodů pro jednotlivé dílčí testy? ANO - NE  
Pokud nesouhlasíte, uveďte proč

---

---

---

**8. Kdyby se zavedla maturitní zkouška v souladu s alternativním katalogem požadavků, jaký by to mělo dopad na výuku anglického jazyka na střední škole? Svoji odpověď zakroužkujte.**

a) pozitivní  
Uveďte proč.

---

---

---

---

---

---

b) negativní  
Uveďte proč.

---

---

---

---

---

---

c) žádný  
Uveďte proč.

---

---

---

---

---

---



Příloha P – 4 Subtest poslechu s porozuměním

**LISTENING PART**

**Student's Sheet**

**1. - 2.** Listen to the recording for tasks 1 and 2. It is a message on an answering machine. Decide which of the answers is correct. Circle the answer a, b, c, or d in your Answer Sheet. Do not write on the Student Sheet. Only one answer is correct.

**(1) Why is Hank calling Bill?**

Because

- a/ he is going to a party.
- b/ he is sick.
- c/ he can't come on time.
- d/ he works little.

**(2) What is the reason why Hank will stop by at Lisa's house?**

- a/ Lisa doesn't feel well.
- b/ He will pick up some food.
- c/ She will help him with his work.
- d/ He will take her with him.

**3. - 6.** Listen to the recording for tasks 3. - 6. It is a dialogue between two friends. Answer the questions below. Write your answers on your Answer Sheet.

**(3) What does the man study?**

**(4) What is the woman majoring in?**

**(5) Why doesn't the woman have to work part-time?**

**(6) Where does the man work part-time?**

**7. - 9.** Listen to the recording for tasks 7 - 9. It is a dialogue between an immigration and customs officer and a traveller. Write the answers on your Answer Sheet. Do not write them in the table below.

The information about the woman

purpose of visit	(7)
place to stay	(8)
things in her luggage	(9)

**10. - 14.** Listen to the recording for tasks 10. - 14. It is a monologue saying something about the life of Isaac Newton, a famous physicist. Answer the questions below. Write the answers on your Answer Sheet. Do not write here.

- (10) In which year was Isaac Newton born?
- (11) Why no one expected him to survive?
- (12) Who did Isaac live with?
- (13) What did he try to measure as a boy?
- (14) When did he become a better student?

15. – 18. Listen to the recording for tasks 15. – 18. It is a dialogue between Carl and Pancho. Fill in the missing information about Pancho in your Answer Sheet.

- Country he comes from: (15) \_\_\_\_\_ .
- Number of sisters: (16) \_\_\_\_\_
- Father's job: (17) \_\_\_\_\_
- Their family store sells (write at least 3 things): (18) \_\_\_\_\_

19. – 23. Listen to the recording for tasks 19. - 23. It is a dialogue between a man and woman. Decide whether the statements below are true or false. If you think they are true write "T" into your Answer Sheet. If you think they are false, write "F" on your Answer Sheet.

- (19) The man and the woman will visit art museum.
- (20) The Indian restaurant is close to the place they are going to visit.
- (21) The woman would like to see unique animals in the zoo.
- (22) They both would like to go shopping.
- (23) The man and woman will go for a walk at the seashore.

24. – 30. Listen to the recording for tasks 24. – 30. The recording is from the Voice of America Special English and it is about the American city of Charleston. Fill in the missing information on your Answer Sheet.

- Charleston is a beautiful city in the south of the United States. Its history is very (24) \_\_\_\_\_. It was first in many different things. For example, there was the first continual (25) \_\_\_\_\_ service in the USA. Among other things, in the American Civil war, the first (26) \_\_\_\_\_ happened in its port.
- Because of the temperature, do not plan your visit in Charleston in (27) \_\_\_\_\_.
- Slaves worked in (28) \_\_\_\_\_ and grew indigo.
- Spanish explorers arrived to the area of Charleston in (29) \_\_\_\_\_.
- Charles Town was founded by the (30) \_\_\_\_\_ settlers in 1670.



# LISTENING PART

## ANSWER SHEET

Name:

School:

1	a b c d
2	a b c d
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	T - F
20	T - F
21	T - F
22	T - F
23	T - F
24	
25	
26	
27	
28	
29	
30	

## LISTENING PART

### ANSWER KEY

1	c
2	a
3	hotel management
4	computer science
5	(She received a four-year) academic scholarship.
6	at a restaurant (near campus)
7	(attend) a teaching convention, touring the capital ( <i>any of the two items is enough as a correct answer</i> )
8	a hotel (downtown)
9	personal belongings/ clothes, books, CD player ( <i>any two of the items are considered to be correct answer</i> )
10	in 1642
11	(Because) he was born early/(because) he was a small baby/ very weak. ( <i>Any of the answers acceptable</i> )
12	(with his) grandmother
13	the speed/strength of the wind
14	after a boy kicked him (in the stomach)
15	Peru
16	six
17	taxi driver
18	bread, eggs, soft drinks, rice, sugar, cookies ( <i>any three of the items are enough</i> )
19	T
20	T
21	F
22	F
23	T
24	rich
25	train
26	battle
27	summer
28	plantations/farms/huge farms
29	152'
30	English

## LISTENING PART

### Tapescript

1. - 2. Listen to the recording for tasks 1 and 2. It is a message on an answering machine. Decide which of the answers is correct. Circle the answer a, b, c, or d in your Answer Sheet. Do not write on the Student Sheet. Only one answer is correct.

**Bill:** Hey, this is Bill. I'm sorry I'm not in. Just leave a message. [*Beep*].

**Hank:** Hey Bill. This is Hank. I'm just calling to let you know that I'll be a little late to the game tomorrow night. I have to work a few extra hours at work to finish a report. I should finish it sometime between seven and eight though. Oh, then I'm planning on dropping by Lisa's house for about an hour since she's been sick recently. And, uh, one more thing. I'll stop at my house to pick up some food for the game. See you then.

3. - 6. Listen to the recording for tasks 3. - 6. It is a dialogue between two friends. Answer the questions below. Write your answers on your Answer Sheet.

**Maria:** Oh, hi Dave. Long time, no see!

**Dave:** Hi Maria. I was in the neighbourhood, so I thought I'd drop by.

**Maria:** Come on in. [Thanks.] Take a seat. Would you like anything to drink? I have Sprite or orange juice.

**Dave:** Sprite would be fine. Uh, so, how have you been?

**Maria:** Oh, not bad. And you?

**Dave:** Oh, I'm doing okay, but school has been really hectic these days, and I haven't had time to relax.

**Maria:** By the way, what's your major anyway?

**Dave:** Hotel management.

**Maria:** Well, what do you want to do once you graduate?

**Dave:** Uh... I haven't decided for sure, but I think I'd like to work for a hotel or travel agency in this area. How about you?

**Maria:** Well, when I first started college, I wanted to major in French, but I realized I might have a hard time finding a job using the language, so I changed majors to computer science. [Oh]. With the right skills, finding a job in the computer industry shouldn't be as difficult.

**Dave:** So, do you have a part-time job to support yourself through school?

**Maria:** Well, fortunately for me, I received a four-year academic scholarship [Wow] that pays for all of my tuition and books.

**Dave:** Wow. That's great.

**Maria:** Yeah. How about you? Are you working your way through school?

**Dave:** Yeah. I work three times a week at a restaurant near campus.

**Maria:** Oh. What do you do there?

**Dave:** I'm a cook.

**Maria:** How do you like your job?

**Dave:** It's okay. The other workers are friendly, and the pay isn't bad.

7. – 9. Listen to the recording for tasks 7 - 9. It is a dialogue between an immigration and customs officer and a traveller. Write the answers on your Answer Sheet. Do not write them in the table below.

**Customs Officer:** Next. Uh, your passport please.

**Woman:** Okay.

**Customs Officer:** Uh, what is the purpose of your visit?

**Woman:** I'm here to attend a teaching convention for the first part of my trip, and then I plan on touring the capital for a few days.

**Customs Officer:** And where will you be staying?

**Woman:** I'll be staying in a room at a hotel downtown for the entire week.

**Customs Officer:** And uh, what do you have in your luggage?

**Woman:** Uh, well, just, just my personal belongings um, . . . clothes, a few books, and a CD player.

**Customs Officer:** Okay. Uh, please open your bag.

**Woman:** Sure.

**Customs Officer:** Okay . . . Everything's fine. Great. Uh, by the way, is this your first visit to the country?

**Woman:** Well, yes and no. Actually, I was born here when my parents were working in the capital many years ago, but this is my first trip back since then.

**Customs Officer:** Well, enjoy your trip.

**Woman:** Thanks.

10. – 14. Listen to the recording for tasks 10. - 14. It is a monologue saying something about the life of Isaac Newton, a famous physicist. Answer the questions below. Write the answers on your Answer Sheet. Do not write here.

Isaac Newton was born in Woolsthorpe, England, on December twenty-fifth, sixteen forty-two. He was born early. He was a small baby and very weak. No one expected him to survive. But he surprised everyone. He had one of the most powerful minds in history. And he lived until he was eighty-four.

Newton's father died before he was born. His mother married again a few years later. She left Isaac with his grandmother.

The boy was not a good student. Yet he liked to make things, such as kites and clocks and simple machines.

Newton also enjoyed finding new ways to answer questions or solve problems. As a boy, for example, he decided to find a way to measure the speed of the wind. On a windy day, he measured how far he could jump with the wind at his back. Then he measured how far he could jump with the wind in his face. From the difference between the two jumps, he made his own measure of the strength of the wind.

Strangely, Newton became a much better student after a boy kicked him in the stomach.

The boy was one of the best students in the school. Newton decided to get even by getting higher marks than the boy who kicked him. In a short time, Newton became the top student at the school.

15. – 18. Listen to the recording for tasks 15. – 18. It is a dialogue between Carl and Pancho. Fill in the missing information about Pancho in your Answer Sheet.

**Carl:** Hi. My name's Carl. Nice to meet you.

**Pancho:** Nice to meet you, too. My name is Francisco.

**Carl:** What?

**Pancho:** Francisco, but all my friends and family back in Peru call me Pancho.

**Carl:** Okay, Pancho. So, tell me about your family.

**Pancho:** Well, I have seven brothers and six sisters.

**Carl:** Wow. That is a big family. So are you the oldest, Pancho?

**Pancho:** No. I'm the second oldest in my family.

**Carl:** So, what do your parents do?

**Pancho:** My father is a taxi driver in Lima, Peru. It's a hard job, but he works hard to support the family.

**Carl:** How about your mother?

**Pancho:** She helps run a small family store with some of my older brothers and sisters.

**Carl:** What kind of store?

**Pancho:** We mainly sell food, like bread, eggs, soft drinks, rice, sugar, and cookies. Things that people buy every day.

19. – 23. Listen to the recording for tasks 19. - 23. It is a dialogue between a man and woman. Decide whether the statements below are true or false. If you think they are true, circle "T" in your Answer Sheet. If you think they are false, circle "F" in your Answer Sheet.

**Man:** So, what do you want to do tomorrow?

**Woman:** Well, let's look at this city guide here. Uh, here's something interesting. Why don't we first visit the art museum in the morning?

**Man:** Okay. I like that idea. And where do you want to eat lunch?

**Woman:** How about going to an Indian restaurant? The guide recommends one downtown a few blocks from the museum.

**Man:** Now that sounds great. After that, what do you think about visiting the zoo? Well, it says here that there are some very unique animals not found anywhere else.

**Woman:** Well, to tell the truth, I'm not really interested in going there. Yeah. Why don't we go shopping instead? There are supposed to be some really nice places to pick up souvenirs.

**Man:** Nah, I don't think that's a good idea. We only have few travellers checks left, and I only have fifty dollars left in cash.

**Woman:** No problem. We can use YOUR credit card to pay for MY new clothes.

**Man:** Oh, no. I remember the last time you used MY credit card for YOUR purchases.

**Woman:** Oh well. Let's take the subway down to the seashore and walk along the beach.

**Man:** Now that sounds like a wonderful plan.

**24. – 30.** Listen to the recording for tasks 24. – 30. The recording is from the Voice of America Special English and it is about the American city of Charleston. Fill in the missing information on your Answer Sheet.

Welcome to THIS IS AMERICA in VOA Special English. I'm Steve Ember with Shirley Griffith. This week, come along to one of the most beautiful and historic cities in the United States – Charleston, South Carolina.

Charleston has a very rich history. It is the only city in the United States that has defended itself from American Indians, fierce pirates, Spanish ships, French soldiers, and British forces. It was first in many things. Charleston had the first continual train service in the United States.

It built the first museum and the first public flower garden in America. And the first battle of the American Civil War took place on a very small but important island in its port.

Plan your visit to Charleston for early spring, late autumn or the winter months. The citizens of Charleston will tell you their lovely city is not fun in the summer. It is extremely hot. The summer heat is important to the history of Charleston.

Early settlers owned huge farms called plantations. Many plantation owners forced slaves to do the work needed to grow indigo in the extreme heat. Slavery became important to the economy of Charleston. The plantations, indigo and slavery are part of the history of the city.

At least three Indian tribes were living in the area that became Charleston when Spanish explorers arrived in fifteen twenty one. The Spanish explorers, and later French explorers tried to establish settlements near that area but none lasted.

English settlers first came to the area in sixteen seventy. They established a town. They called it Charles Town in honor of the English King, Charles the Second. The people of the city changed its name to Charleston in seventeen eighty-three...

Příloha P – 5 Subtest čtení s porozuměním

**READING PART**  
**Student Sheet**

*Time limit: 45 minutes*

Instructions

*Read the following four short texts A, B, C, and D. Then read the three questions below the text and answer them by circling a, b, c, or d in your Answer Sheet. Do not write your answers here. Only one answer is correct. There are four texts, but only three questions.*

A) The Texas Folklife Festival is open 5 – 11 p.m. Aug. 5 – 6; noon – 11 p.m. Aug. 7; noon – 9 p.m. Aug. 8. Admission is for the entire day. If you wish to leave the grounds and return later in the day, please be sure to have your hand stamped at the gate.

The festival's 11 stages are located throughout the grounds and feature all types of ethnic and cultural entertainment. For a complete list of performance times, see the daily schedules. Admission is \$8; \$2 for children 6 – 12; free for children 5 and younger.

B) A downtown bookstore with an outstanding selection in all categories. Specialties include Texas and regional books, as well as South-West and regional cookbooks.

- Full service with knowledgeable staff
- Large in-stock selection of books, cards, calendars, posters and other merchandise
- Free special order service
- Free gift wrap
- Toll-free ordering: 1-800-688-3927
- Shipping available to any and all destinations
- Free delivery to downtown hotels

C) How to work with the meters

- Look for the space number on the street and the meter number on the sign nearby.
- Press any button to start.
- Enter space number.
- Deposit coins or \$1 bills.
- Press receipt button.

D) Upon arrival you will be assigned a room in the dormitory. You should inspect your room as soon as possible. After the inspection, you will be asked to sign a room inventory sheet. If there are any problems, make sure to write them on this sheet. Remember, after the first week you become responsible for replacing items that are damaged. It is also wise to inform your senior student of any problems and have him/her present when you inspect your room and when you report any problem.

- (1) Which of the texts will you probably read after parking your car downtown?
- (2) Which of the texts gives the information about student accommodation?
- (3) Which of the texts informs the customers about various kinds of goods and services?

Instructions

Read the text. Then read the three sentences below it and fill in the gaps according to the content of the text. Write your answers into your **Answer Sheet**.

English Language Center, Language Institute, San Antonio, Texas  
American Members of International Goodwill to Others (AMIGO)  
Sponsorship Program

Welcome to the Language Institute, English Language Center. AMIGO stands for "American Members of International Goodwill to Others". AMIGO is also a Spanish word for "friend". The AMIGO Sponsorship Program is designed to introduce you to Americans providing you an opportunity to understand our American culture, our human rights values, and to give you an opportunity to practice your English language skills. We realize it is difficult for you to be in a foreign country without your families for a long time. However, the head of the institute has requested that Americans and students in our AMIGO Program do not date each other. This program is not a dating service.

- The word "amigo" has \_\_\_\_\_ (4) meanings.
- One of the purposes of the AMIGO program is to help the students to \_\_\_\_\_ (5) their English.
- Students and Americans should not \_\_\_\_\_ (6).

Instructions

The following three paragraphs are about things that can be harmful for your health. Below the three paragraphs there are five headings. For each paragraph, choose one heading which suits the content of the paragraph. Circle your answers a, b, c, d, or e in your **Answer Sheet**.

(7) \_\_\_\_\_  
Colas, hot chocolate, tea and coffee are loaded with caffeine which stimulates the nervous system. Too much caffeine can overstimulate the brain and cause irritation, faster heartbeat and increased blood pressure. Limit your intake of caffeinated drinks to no more than two per day. Have them early, rather than late, so they don't interfere with your sleep.

(8) \_\_\_\_\_  
When you feel an energy low, the first thing you reach for is sugar: a candy bar, chocolate, cookies or pastry. That's a no-no. Refined sugar products give you an initial rush, followed by an energy comedown. When you feel low, try fruit juice, fresh fruit, vegetable soup or yoghurt. The carbohydrates will keep you going until meal time.



(9) \_\_\_\_\_

25 percent of all Americans doesn't eat before going to work. But a morning meal is absolutely essential to supply you with day-time energy reserves. A light, low-fat breakfast stimulates the metabolism, helps ensure that you get an even supply of calories and nutrients and saves digestive energy.

- a) Have enough drinks
- b) Feeling depressed
- c) Between meals
- d) Just two a day
- e) Skipping breakfast

Instructions

*Read the instructions on treating and preventing influenza (flu). Then read the statements below the test and decide whether they are TRUE or FALSE. In your Answer Sheet circle "T" if you think the statement is true or "F" if you think it is false.*

How is influenza treated?

Two prescription drugs are available to treat the symptoms of some types of flu. These drugs are helpful if you start taking them within the first 48 hours of illness. They can make flu symptoms less severe. Also, the symptoms may not last as long. These drugs need to be taken only 3 to 5 days. A common side effect of the drug is dizziness.

It is very important to drink a lot of liquids. Water, juice, and non-caffeinated drinks are best. Especially, when you have a high fever, your body needs much more liquid than when you are healthy.

Many health care providers recommend using non-prescriptions medicines to lower a fever. But some doctors think you should not use aspirin if you have the flu. They say fever is part of the immune system's reaction to infection, so it's better to let a fever run its course than to try to lower it. Letting the fever run its course, however, can be dangerous in children and elderly. Also, most healthy adults feel much better if the fever is decreased even 1 or 2 degrees.

What can I do to prevent influenza?

The best way to prevent influenza is to get a flu shot every October. Flu shots are about 70% effective in preventing influenza. Very few people have side effects. About 1 – 2% may have mild flu symptoms in the first 24 hours after getting a flu shot. People who are allergic to eggs should not get flu shots.

- (10) Prescription drugs can be taken only for 48 hours.
- (11) It is always necessary to lower a fever during flu.
- (12) Flu vaccines help most people to prevent flu.
- (13) It is enough to get a flu shot once a year
- (14) All people are recommended to get flu shots.

Instructions

Read the text that is an extract from a narration of a city tour guide. Read the questions below and circle the answer a, b, c, or d in your Answer Sheet.

Today, we will play tourist and introduce you to the life style and traditions of the city of San Antonio. This tour is designed to acquaint you with place and history that makes this city one of the four unique cities of the USA.

The style of San Antonio was born of the traditions of Indian tribes, as well as of Mexican domination for nearly 150 years. With a population over a million people and a mixture of races and cultures, there are surprises everywhere.

Our first stop is located in our largest city park. The Japanese Tea Gardens is a man-made garden designed from an original limestone quarry. This was once the site of the largest cement plant west of the Mississippi river. This land was donated to the city and the park manager, and a Japanese artist designed this beautiful setting. In the gardens, you can see hundreds of varieties of plants on either side of the path that passes ponds and fountains.

From this park, we will go up into the Tower of Americas. Completed in 1968, the tower stands as a reminder of HemisFair, Texas' world's fair for which it was built. The tower is still one of the tallest freestanding structures in the western hemisphere. Atop the 750-foot tower you can truly see San Antonio. On a clear day, the entire city is in view.....

(15) The city of San Antonio

- a) was also influenced by Mexican and Indian culture.
- b) is a mixture of Mexican and American culture.
- c) was inhabited by Mexicans for more than 150 years.
- d) is an example of a typical American city.

(16) The Japanese Tea Gardens

- a) belong to a Japanese artist.
- b) belong to the city.
- c) are to the east from the Mississippi River.
- d) are next to the city park.

(17) The Tower of Americas

- a) stands as a lookout tower.
- b) will be part of American fair .
- c) is the tallest in the world.
- d) dominates the city park

Instructions

Read the leaflet from the International Institute that is a part of the University of Salford, UK. Answer the questions below the text by circling a, b, c, or d in your Answer Sheet.

University of Salford  
International Institute  
SUMMER GENERAL AND ACADEMIC  
ENGLISH COURSES, 2005

**A)** 29 June to 11 September 2005  
11 weeks duration  
Fee: £ 1 650

This course will begin with an introductory four-week block of General English with options in English for Academic Purposes. The course will then progress into an Academic English Course with two further entry points as detailed below.

**B)** 29 June to 24 July 2005  
4 weeks duration  
Fee: £ 640

As a free standing block this is essentially a General English Course and such is suitable for those students who are not necessarily going to continue with studies in the UK or at home. This block would also be suitable for those students continuing to an Academic English Course.

**C)** 27 July to 14 August 2005  
3 weeks duration  
Fee: £ 480

As a three-week block of Academic English this course will be of particular benefit to those international students who hope to continue their studies at the University of Salford. It would also be useful for students intending to study at other universities in the UK or in their own countries.

**D)** 17 August to 11 September 2005  
4 weeks duration  
Fee: £ 640

As a four-week block of intensive study of Academic English this course will be of particular benefit to those international students who hope to continue their studies at the University of Salford. It would also be useful for students intending to study at other universities in the UK or in their own countries.

The above prices are inclusive of a full social programme. University accommodation can be arranged (at extra cost) for participants in the above courses.

- (18) Which course is good for a student who does not want to go on studying?  
(19) Which course is for students who cannot go to Britain in August?  
(20) Which course is General English only?  
(21) Which course would be appropriate for a student who would like to study Academic English and has got only £ 500?

Instructions

Read the following text with some tips on how to stop smoking. Answer the questions below. Write the answers *into your Answer Sheet*.

Help yourself to quit

- Stop day – choose a day to stop and stop completely on that day.
- Tell everyone you're stopping - this will give you support to stick with it when you feel like a cigarette.
- Ash cash – work out how much you will save and plan how you will spend it.
- Think positively – you're a non-smoker, not a smoker who's given up. When you're offered a cigarette, say 'no thanks, I don't smoke.' - every time.
- Not everyone succeeds the first time. It's a strong addiction, but don't give up trying.
- Ask you family doctor about the availability of techniques like nicotine chewing gum, stop-smoking groups, hypnosis or acupuncture therapy.

- (22) What should you do to avoid having a cigarette?  
(23) What is ash cash?  
(24) What can you do to have your mouth busy?

Instructions

Read the following text that is a Pharmacy Boots leaflet. According to the content of the leaflet, fill in the gaps of the summary. You will need just one word. Write the answers *into your Answer Sheet*.

At your service

Pharmacy Boots  
28/29 great Western Arcade,  
Birmingham, B2 5HU  
Phone: 021 -233 3538

#### Help and advice

Our trained staff will be pleased to offer help and advice on your prescribed medicine or any other healthcare matters.

#### Patient medication records

If you join our Patient Medication Record Scheme, we can keep a confidential record of the medicines given to you. This service is free. Once you have registered, we will give you a record card with your personal number. Show your card the next time you bring in a prescription and we will be able to serve you more quickly. We can check that you can safely take this medicine and that it is right for you.

#### Unwanted medicines

To avoid accidents at home, we are happy to safely destroy unwanted medicine for you. Just bring them here.

#### Complaints

If you think we can improve our services provided by this pharmacy, please tell us. The local Family Health Services Authority can also help.

#### Urgent prescriptions

If your doctor gives you an urgent prescription that needs to be dispensed outside our normal opening hours, please contact your local police station. They will arrange for a pharmacist to be called out.

The staff of pharmacy *Boots* is \_\_\_\_\_ (25) and they will give you \_\_\_\_\_ (26) on the medicine you need. Patient Medication record Scheme is free of charge and it will help the pharmacy staff to serve you more quickly. They can also decide whether it is \_\_\_\_\_ (27) for you to take the medicine. The pharmacists will be also pleased to destroy the medicine that you don't \_\_\_\_\_ (28). If you have any \_\_\_\_\_ (29), you can contact your local Family Health Services Authority. It may happen that you will need a medicine urgently and outside opening hours of the pharmacy. In such a case, the \_\_\_\_\_ (30) will call out the pharmacist.

**READING PART**

**ANSWER SHEET**

*Name:*

*School:*

1	a b c d
2	a b c d
3	a b c d
4	
5	
6	
7	a b c d e
8	a b c d e
9	a b c d e
10	T - F
11	T - F
12	T - F
13	T - F
14	T - F
15	a b c d
16	a b c d
17	a b c d
18	a b c d
19	a b c d
20	a b c d
21	a b c d
22	
23	
24	
25	
26	
27	
28	
29	
30	

## READING PART

### ANSWER KEY

1	c
2	d
3	b
4	two/2
5	practise/improve
6	date/date each other
7	d
8	c
9	e
10	F
11	F
12	T
13	T
14	F
15	a
16	b
17	a
18	b
19	b
20	b
21	c
22	tell everyone you are stopping
23	the money you save when you don't smoke
24	have a (nicotine) chewing gum
25	trained/qualified
26	advice
27	safe/right
28	want/need
29	complaint(s)
30	police

## WRITING PART

### Student's Sheet

*Time limit: 35 minutes*

*Your writing will be assessed for the content, vocabulary and grammar. Each point should be mentioned.*

#### Task 1

*You are on a trip with your schoolmates in the UK. You are staying with an English family. They are at work now and you would like to go somewhere with someone. Write a short message for the family you are staying with and inform them about the following things:*

- the place you are going to
- the person you are going with
- the reason
- the time of coming back
- no food in the fridge

#### Task 2

*You would like to work as an au-pair in Great Britain. An agency has found a family for you. Now the family wants you to write them a letter about yourself in which you should give details about the following points:*

- family background (parents, brothers, sisters)
- education
- hobbies and interests
- reasons why you want to do the job
- experience with taking care of children.

#### Task 3

*A foreign friend of yours would like to come for holiday to the Czech Republic. He/she would like to spend skiing/cycling holiday here. Write an email to him/her in which you will give him/her information he/she has asked you for. You can also express your own opinion/experience. Here are the points you should mention:*

- the place/places to visit and why;
- the kind of accommodation to stay at and why;
- the weather during the holiday;
- the recommended means of transportation in the country;
- the Czech food.



## **WRITING PART**

### **Answer Sheet**

*Name:*

*School:*

*Task 1*

*Task 2*

**Task 3**

**WRITING PART**

Write a letter

Task 1

You have been invited to a party at your friend's house. Write a letter to your friend accepting the invitation. In your letter, mention the date, time and place of the party. Also, mention what you will wear and what you will bring to the party.

Task 2

You have been invited to a party at your friend's house. Write a letter to your friend declining the invitation. In your letter, mention the date, time and place of the party. Also, mention why you cannot attend and what you would like to do instead.

You have been invited to a party at your friend's house. Write a letter to your friend asking for more information about the party. In your letter, mention the date, time and place of the party. Also, mention what you would like to know about the party.

You have been invited to a party at your friend's house. Write a letter to your friend thanking them for the invitation. In your letter, mention the date, time and place of the party. Also, mention how much you enjoyed the party and what you will do to reciprocate.

Write a letter

- to a friend
- to a family member
- to a teacher
- to a boss
- to a friend who is moving
- to a friend who is getting married

Příloha P – 7 Subtest mluvení

**SPEAKING PART**

**Student's Sheet**

*Task 1*

*Tell us something about your family background (parents – their jobs, brothers and sisters, the place you live in...). Be prepared to answer some questions from the examiners.*

*Task 2*

*This task is a role play between you and your examiner. First read the role. If you don't understand anything, ask your examiner.*

An English-speaking tourist has met you on the street. He/she wants to get to the city/town centre. Give him/her instructions and some more information that he/she will ask you for. The examiner will play the role of the foreign tourist.

*Task 3*

*In this task, you will ask your examiner some questions. You can take notes about his/her answers, because then you will report them to the other examiner. Here is the list of the questions in Czech. You will ask them in English.*

Zeptejte se zkoušející(ho):

- co dělal(a) včera večer;
- co měl(a) k večeři;
- zda je večer radši doma nebo mimo domov;
- jaká je jeho/její oblíbená restaurace a proč;
- co bude dělat příští víkend.

## Teacher's Sheet

*At the beginning of the oral test, introduce your colleague and yourself. Try to create a relaxed and friendly atmosphere and to put the candidates at ease. Do not correct their mistakes, do not take notes. Ask clear questions.*

### Task 1

*Ask the candidate to tell you something about his/her family background (parents – their jobs, brothers and sisters, the place they live in...). Don't let them talk a pre-learned monologue; ask questions.*

### Task 2

*This is a role play between you and the candidate.*

An English-speaking tourist (you) have met a Czech (the candidate) on the street. You want to get to the city/town centre. The candidate will give you instructions and some more information that you will ask for. (how to get there, which means of transportation to use, how long it takes, what you can see there, what is interesting to visit...).

Possible questions to ask the candidate:

- Could you tell me how to get to the city centre?/ Excuse me, how can I get to the city centre?
- How long does it take to get there?
- Where can I buy a tram/bus ticket?
- Do you know the price?
- Is there a restaurant/café/art gallery/museum in the city centre? How can I get there?

### Task 3

*In this task, you will be asked some questions. The candidate is allowed to take notes about your answers, because then he/she will report them to the other examiner. Answer the questions clearly, but naturally. If the candidate does not seem to understand, rephrase them.*

Possible answers:

- I was out at a party last night.
- I had different sandwiches, barbecue, fruits and dessert.
- I prefer going out, sometimes I go to a restaurant but mostly I like to go swimming.
- My favourite restaurant is a pizza restaurant in my neighbourhood. The food is excellent there.
- I'm going to make a trip to the mountains with my family because we all like skiing.

## Příloha P – 8 Hodnotící kritéria písemného projevu

### Hodnotící kritéria písemného projevu

#### Examiner's Sheet

#### Úkol 1

##### Hodnocení

- *Maximální počet bodů za tento úkol je 5. Každý bod odpovídá jedné obsahové položce specifikované v zadání.*
- *Pokud obsahová položka splní cíl komunikace, tedy pokud bude obsah sdělení v souladu se specifikací zadání, dostane kandidát za každou obsahovou položku ve vzkazu 1 bod;*
- *V každé obsahové položce, která se hodnotí 1 bodem, používá kandidát většinou správně vysokofrekvenční slovní zásobu a jednoduché gramatické struktury (alespoň 3 základní časy: budoucí, přítomný, minulý); je přípustné, aby vzkaz obsahoval i méně závažné chyby v pravopise, slovní zásobě nebo gramatice;*
- *Pokud kandidát obsahovou položku do vzkazu nezahrnul, pokud nebyl splněn cíl komunikace, nebo pokud výskyt gramatických a lexikálních chyb vážně narušuje porozumění textu je položka hodnocena 0 body.*

#### Úkol 2

##### Hodnocení

- *Maximální počet bodů za tento úkol je 10.*
- *Pokud obsahová položka splní cíl komunikace, tedy pokud bude obsah sdělení v souladu se specifikací zadání, dostane kandidát za každou položku v dopisu 2 body;*
- *V každé obsahové položce, která se hodnotí 2 body, používá kandidát většinou správně vysokofrekvenční slovní zásobu a jednoduché gramatické struktury (alespoň 3 základní časy: budoucí, přítomný, minulý); je přípustné, aby dopis obsahoval i méně závažné chyby v pravopise, slovní zásobě nebo gramatice;*
- *V každé obsahové položce, která se hodnotí 1 bodem, kandidát splnil cíl komunikace pouze částečně anebo písemný projev obsahuje závažnější chyby v pravopise, slovní zásobě nebo gramatice, které však nenaruší úplně porozumění textu;*

- *Pokud kandidát obsahovou položku do dopisu nezahrnul, pokud nebyl splněn cíl komunikace, nebo pokud výskyt gramatických a lexikálních chyb vážně narušuje porozumění textu je položka hodnocena 0 body.*

### *Úkol 3*

#### *Hodnocení.*

- *Maximální počet bodů za tento úkol je 10.*
- *Pokud obsahová položka splní cíl komunikace, tedy pokud bude obsah sdělení v souladu se specifikací zadání, dostane kandidát za každou položku v dopisu 2 body;*
- *V každé obsahové položce, která se hodnotí 2 body, používá kandidát většinou správně vysokofrekvenční slovní zásobu a jednoduché gramatické struktury (alespoň 3 základní časy: budoucí, přítomný, minulý); je přípustné, aby dopis obsahoval i méně závažné chyby v pravopise, slovní zásobě nebo gramatice;*
- *V každé obsahové položce, která se hodnotí 1 bodem, kandidát splnil cíl komunikace pouze částečně anebo písemný projev obsahuje závažnější chyby v pravopise, slovní zásobě nebo gramatice, které však nenaruší úplně porozumění textu;*
- *Pokud kandidát obsahovou položku do vzkazu nezahrnul, pokud nebyl splněn cíl komunikace, nebo pokud výskyt gramatických a lexikálních chyb vážně narušuje porozumění textu je položka hodnocena 0 body.*

## Příloha P – 9 Hodnotící kritéria ústního projevu

### Hodnotící kritéria ústního projevu

#### Charakteristika jednotlivých testových úloh

##### *1. úloha: krátký úvodní rozhovor*

Kandidát je dotázán na jednoduché, jemu známé téma, např. rodina, záliby, budoucí zaměstnání, volný čas apod. Aby se předešlo odříkávání naučených témat, zkoušející v průběhu plnění úlohy reaguje doplňujícími otázkami, které mají vést u kandidáta k bezprostřední (tj. nenaučené) produkci mluveného slova.

##### *2. úloha: sehrání situační role*

Kandidát sehraje situační roli se zkoušejícím. Situační role vycházejí z každodenního života; nejčastější kontexty: lékař, nakupování, studium, cestování, hotel, restaurace. Kandidát obdrží kartičku se zadáním situační role v anglickém jazyce. Použité situační role simulují autentické situace, ve kterých by se kandidát mohl ocitnout při komunikaci s anglicky mluvícími osobami ať už v domácím nebo cizojazyčném prostředí. Důležité je, aby hrál roli sobě přirozenou (tedy např. roli pacienta a ne lékaře, pokud role simuluje situaci v ordinaci).

##### *3. úloha: získání konkrétních informací*

Kandidát má na kartičce napsáno pět informací, které má zjistit od examinátora. Požadované informace jsou napsány v českém jazyce. Od kandidáta se očekává zjištění požadovaných informací od jednoho examinátora a poté předání těchto informací jinému examinátorovi v anglickém jazyce.

#### Maximální počet bodů:

- 1. úloha: 5 bodů*
- 2. úloha: 10 bodů*
- 3. úloha: 10 bodů*

## Hodnoticí škála

### 1. úloha: max. 5 bodů

#### 4 - 5 bodů

- Kandidát splnil cíl komunikace – podat informaci a zodpovědět otázky související s podanou informací;
- používal většinou správně vysokofrekvenční slovní zásobu a jednoduché gramatické struktury (alespoň 3 základní časy: budoucí, přítomný, minulý);
- je přípustné, aby se v ústním projevu vyskytovaly i méně závažné chyby ve výslovnosti, slovní zásobě a/nebo gramatice;

#### 2 – 3 body

- Kandidát splnil cíl komunikace - podat informaci a zodpovědět otázky související s podanou informací;
- chyby ve výslovnosti, gramatice a/nebo slovní zásobě však mohou způsobit komunikační problémy;

#### 1 bod

- Kandidát nesplnil cíl komunikace - podat informaci a zodpovědět otázky související s podanou informací;
- ojediněle prezentoval částečně srozumitelný vzorek řeči

#### 0 bodů

- Kandidát nesplnil cíl komunikace - podat informaci a zodpovědět otázky související s podanou informací;
- vzorek řeči nebyl prezentován nebo byl zcela nesrozumitelný.

### 2. úloha: max. 10 bodů

#### 8 – 10 bodů

- Kandidát splnil cíl komunikace – sehrál úspěšně svou roli a lze předpokládat, že by v reálném životě zvládl podobnou komunikační situaci bez problémů;
- používá většinou správně vysokofrekvenční slovní zásobu a jednoduché gramatické struktury (alespoň 3 základní časy: budoucí, přítomný, minulý);
- je přípustné, aby projev obsahoval i méně závažné chyby ve výslovnosti, slovní zásobě a/nebo gramatice;

#### 5 – 7 bodů

- Kandidát splnil cíl komunikace – sehrál svou roli uspokojivým způsobem a lze předpokládat, že v reálném životě by podobnou situaci zvládnul s méně závažnými problémy;
- chyby ve výslovnosti, gramatice a/nebo slovní zásobě však mohou způsobit komunikační problémy;



### **3 – 4 body**

- Kandidát nesplnil cíl komunikace – sehrání role mu působilo značné potíže a lze předpokládat, že v reálném životě by v podobné komunikační situaci neuspěl;
- prezentoval však částečně srozumitelný vzorek řeči;

### **0 – 2 body**

- Kandidát nesplnil cíl komunikace - sehrání role mu působilo značné potíže a lze předpokládat, že v reálném životě by v podobné komunikační situaci neuspěl;
- vzorek řeči nebyl prezentován nebo byl zcela nesrozumitelný.

## **3. úloha: max. 10 bodů**

### **8 – 10 bodů**

- Kandidát splnil cíl komunikace – otázky i předání informací byly srozumitelné a lze předpokládat, že by v reálném životě zvládl podobnou komunikační situaci bez problémů;
- používá většinou správně vysokofrekvenční slovní zásobu a jednoduché gramatické struktury (alespoň 3 základní časy: budoucí, přítomný, minulý);
- je přípustné, aby projev obsahoval i méně závažné chyby ve výslovnosti, slovní zásobě a/nebo gramatice;

### **5 – 7 bodů**

- Kandidát splnil cíl komunikace – otázky i předání informací byly většinou srozumitelné a lze předpokládat, že v reálném životě by podobnou situaci zvládnul s menšími problémy;
- chyby ve výslovnosti, gramatice a/nebo slovní zásobě však mohou způsobit komunikační problémy;

### **3 – 4 body**

- Kandidát nesplnil cíl komunikace – otázky i předání informací mu působily značné potíže a lze předpokládat, že v reálném životě by v podobné komunikační situaci neuspěl;
- ojediněle prezentoval srozumitelný vzorek řeči;

### **0 – 2 body**

- Kandidát nesplnil cíl komunikace – nebyl vůbec schopen získat a předat potřebné údaje a lze předpokládat, že v reálném životě by v podobné komunikační situaci neuspěl;
- vzorek řeči nebyl prezentován nebo byl zcela nesrozumitelný.

## Příloha P – 10 Průvodní instrukce pro zadávající a dotazník k subtestu čtení

### Průvodní instrukce

1. Tento test je určen k ověření u žáků středních škol. Je součástí disertační práce.
2. Má testovat dosažení úrovně řečové dovednosti čtení A2 podle Evropského referenčního rámce.
3. Časové limity jsou uvedeny na začátku každé části testu.

#### READING PART

- Žáci zaznamenávají své odpovědi do odpovědního archu (Answer Sheet).
- Testy opravím sama a pošlu jmenný seznam s jejími výsledky.
- Jakékoliv připomínky z Vaší strany nebo ze strany žáků uveďte prosím, do stručného dotazníku READING PART – FEEDBACK.

*Děkuji mnohokrát za Vaši spolupráci.*



## Příloha P – 11 Průvodní instrukce pro zadávající a dotazník k subtestu psaní

### Průvodní instrukce

1. Tento test je určen k ověření u žáků středních škol. Je součástí disertační práce.
2. Má testovat dosažení úrovně řečové dovednosti psaní A2 podle Evropského referenčního rámce.
3. Časové limity jsou uvedeny na začátku každé části testu.

#### WRITING PART

- Žáci píší tuto část testu na *Answer Sheet*.
- Prosím o laskavé ohodnocení této části testu. Kritéria hodnocení přiloženy.
- NEPIŠTE, prosím, POČET BODŮ přímo DO ODPOVĚDNÍCH ARCHŮ.
- Přiložte seznam jmen s bodovým hodnocením. Písemné práce budou hodnoceny ještě další osobou, takže je důležité, aby ji Vaše hodnocení neovlivnilo.
- Jakékoliv připomínky z Vaší strany nebo ze strany žáků uveďte prosím, do stručného dotazníku WRITING PART – FEEDBACK.





## Conventional Item and Test Analysis Program

Item analysis for data from file C:\Program Files\ITEMAN\Ria-disertace.dat

Date: 26 V 2006

Time: 7,53

Seq. No. Key	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	High	Point Biser.
1	1-1	.75	.45	.46	A	.11	.10	.00	-.13
					B	.12	.32	.03	-.44
					C	.75	.52	.97	.46 *
					D	.02	.06	.00	-.12
					Other	.00	.00	.00	
2	1-2	.88	.32	.38	A	.05	.13	.00	-.23
					B	.01	.03	.00	-.08
					C	.05	.13	.00	-.21
					D	.88	.68	1.00	.38 *
					Other	.01	.00	.00	-.21
3	1-3	.87	.32	.40	A	.05	.13	.00	-.19
					B	.87	.68	1.00	.40 *
					C	.04	.10	.00	-.23
					D	.02	.06	.00	-.18
					Other	.02	.00	.00	-.18
4	1-4	.44	.60	.54	1	.44	.16	.76	.54 *
					2	.46	.52	.24	-.29
					Other	.10	.00	.00	-.40
5	1-5	.82	.39	.43	1	.82	.58	.97	.43 *
					2	.14	.29	.03	-.30
					Other	.05	.00	.00	-.29
6	1-6	.88	.26	.39	1	.88	.74	1.00	.39 *
					2	.04	.10	.00	-.25
					Other	.08	.00	.00	-.29
7	1-7	.74	.36	.33	A	.17	.32	.09	-.22
					B	.09	.16	.03	-.22
					C	.00	.00	.00	
					D	.74	.52	.88	.33 *
					E	.00	.00	.00	
					Other	.00	.00	.00	

8	1-8	.48	.28	.26	A	.19	.35	.09	-.29
					B	.31	.29	.30	-.03
					C	.48	.32	.61	.26 *
					D	.01	.00	.00	.06
					E	.00	.00	.00	
					Other	.01	.00	.00	-.08
9	1-9	.98	.06	.22	A	.00	.00	.00	
					B	.00	.00	.00	
					C	.02	.06	.00	-.22
					D	.00	.00	.00	
					E	.98	.94	1.00	.22 *
					Other	.00	.00	.00	
10	1-10	.94	.16	.31	1	.05	.13	.00	-.24
					2	.94	.84	1.00	.31 *
					Other	.01	.00	.00	-.23
11	1-11	.79	.42	.38	1	.21	.45	.03	-.38
					2	.79	.55	.97	.38 *
					Other	.00	.00	.00	
12	1-12	.91	.16	.32	1	.91	.81	.97	.32 *
					2	.08	.16	.03	-.26
					Other	.01	.00	.00	-.23
13	1-13	.89	.23	.24	1	.89	.74	.97	.24 *
					2	.11	.26	.03	-.24
					Other	.00	.00	.00	
14	1-14	.92	.19	.33	1	.06	.13	.00	-.24
					2	.92	.81	1.00	.33 *
					Other	.02	.00	.00	-.23
15	1-15	.64	.37	.31	A	.64	.45	.82	.31 *
					B	.13	.19	.09	-.17
					C	.21	.32	.06	-.21
					D	.02	.03	.00	-.09
					Other	.01	.00	.00	.08
					A	.16	.29	.00	-.31
16	1-16	.70	.68	.58	B	.70	.32	1.00	.58 *
					C	.07	.23	.00	-.35
					D	.04	.10	.00	-.18
					Other	.03	.00	.00	-.18
					A	.61	.61	.82	.14 *
					B	.16	.16	.06	-.08
17	1-17	.61	.21	.14	C	.03	.10	.00	-.24
					D	.20	.13	.12	.00
					Other	.00	.00	.00	
					A	.14	.19	.06	-.14
					B	.65	.42	.85	.36 *
18	1-18	.65	.43	.36	C	.04	.10	.00	-.20
					D	.15	.26	.09	-.20
					Other	.02	.00	.00	-.11
					A	.14	.19	.06	-.14
					B	.65	.42	.85	.36 *



19	1-19	.62	.40	.39	A	.11	.26	.09	-.24	
					B	.62	.39	.79	.39	*
					C	.05	.06	.03	-.09	
					D	.21	.26	.09	-.21	
					Other	.02	.00	.00	-.07	
20	1-20	.66	.46	.39	A	.26	.39	.03	-.35	
					B	.66	.48	.94	.39	*
					C	.00	.00	.00		
					D	.05	.10	.03	-.08	
					Other	.02	.00	.00	-.09	
21	1-21	.95	.06	.17	A	.00	.00	.00		
					B	.00	.00	.00		
					C	.95	.94	1.00	.17	*
					D	.04	.03	.00	-.13	
					Other	.01	.00	.00	-.12	
22	1-22	.64	.46	.41	1	.64	.39	.85	.41	*
					2	.12	.13	.12	.02	
					Other	.25	.00	.00	-.48	
23	1-23	.40	.75	.61	1	.40	.10	.85	.61	*
					2	.39	.45	.15	-.23	
					Other	.21	.00	.00	-.46	
24	1-24	.74	.52	.52	1	.74	.45	.97	.52	*
					2	.09	.06	.03	-.15	
					Other	.17	.00	.00	-.48	
25	1-25	.34	.79	.66	1	.34	.03	.82	.66	*
					2	.35	.23	.18	-.02	
					Other	.31	.00	.00	-.66	
26	1-26	.42	.54	.42	1	.42	.16	.70	.42	*
					2	.35	.23	.30	.10	
					Other	.24	.00	.00	-.60	
27	1-27	.51	.81	.67	1	.51	.06	.88	.67	*
					2	.16	.19	.12	-.09	
					Other	.33	.00	.00	-.64	
28	1-28	.56	.75	.63	1	.56	.13	.88	.63	*
					2	.13	.06	.09	-.02	
					Other	.31	.00	.00	-.67	
29	1-29	.45	.75	.59	1	.45	.06	.82	.59	*
					2	.23	.13	.15	.04	
					Other	.32	.00	.00	-.67	
30	1-30	.55	.87	.70	1	.55	.13	1.00	.70	*
					2	.11	.06	.00	-.07	
					Other	.35	.00	.00	-.69	

There were 110 examinees in the data file.

Scale Statistics

-----  
Scale:                   1  
-----  
N of Items               30  
N of Examinees         110  
Mean                    20.727  
Variance                29.198  
Std. Dev.               5.404  
Skew                    -0.137  
Kurtosis                -0.975  
Minimum                 8.000  
Maximum                30.000  
Median                  20.000  
Alpha                   0.844  
SEM                     2.133  
Mean P                  0.691  
Mean Item-Tot.         0.418  
Mean Biserial          0.593  
Max Score (Low)        16  
N (Low Group)           31  
Min Score (High)       25  
N (High Group)          33

SCALE # 1

Score Distribution Table

Number Correct	Freq- uency	Cum Freq	PR	PCT	
. . . No examinees below this score . . .					
7	0	0	1	0	
8	1	1	1	1	#
9	1	2	2	1	#
10	0	2	2	0	+
11	2	4	4	2	##
12	1	5	5	1	#
13	7	12	11	6	#####
14	3	15	14	3	###
15	6	21	19	5	+#####
16	10	31	28	9	#####
17	3	34	31	3	###
18	10	44	40	9	#####
19	3	47	43	3	###
20	8	55	50	7	+#####
21	2	57	52	2	##
22	6	63	57	5	#####
23	5	68	62	5	#####
24	9	77	70	8	#####
25	6	83	75	5	+#####
26	10	93	85	9	#####
27	5	98	89	5	#####
28	3	101	92	3	###
29	6	107	97	5	#####
30	3	110	99	3	+###

|-----+-----+-----+-----+  
5 10 15 20 25  
Percentage of Examinees

Elapsed Time: 0.156 second

## Literatura

- ALDERSON, J.C. *Assessing Reading*. Cambridge University Press 2000. ISBN 0 521 59999 7
- ALDERSON, J.C.; NAGY, E.; OVEGES, E. *English Language Education in Hungary, Part II*. Budapest : The British Council Hungary 2000.
- ALDERSON, J.C.; CLAPHAM, C.; WALL, D. *Language Test Construction and Evaluation*. 2<sup>nd</sup> ed. Cambridge University Press 1995. ISBN 0 521 47255 5
- ANDERSON, L.W.; KRATHWOHL, D.R. *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Addison Wesley Longman, Inc. 2001. ISBN-0-8013-1903-X
- BACHMAN, L.F. *Fundamental Considerations in Language Testing*. 4<sup>th</sup> ed. Oxford University Press 1997. ISBN 0 19 437003 8
- BACHMAN, L.F.; PALMER, A.S. *Language Testing in Practice*. Oxford University Press 1996. ISBN 0-19-437148-4
- BLOOM, B.S.; ENGELHART, M.D.; FURST, E.J. aj. *Taxonomy of Educational Objectives : The Classification of Educational goals*. 15<sup>th</sup> ed.: David McKay Company, Inc., New York 1969
- BUCK, G. *Assessing Listening*. 1<sup>st</sup> ed. Cambridge University Press 2001. ISBN 0 521 66661 9
- BYČKOVSKÝ, P. *Projektování, realizace a hodnocení kursů pedagogicko-psychologické přípravy pro učitele odborných předmětů na vysokých školách technických*. In: Kotásek, J. (ed.). *Zvyšování kvality pedagogické kvalifikace učitelů vysokých škol*. Praha: Ústav rozvoje vysokých škol 1989, s. 7-36.
- BYČKOVSKÝ, P. Vymezení výukových cílů – předpoklad řízení výuky. *Vysoká škola*, vol. 28, 1980, č. 3, s. 120-133.
- BYČKOVSKÝ, P., ZVÁRA, K. *Konstrukce a analýza testů pro přijímací řízení*. rukopis *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press 2001. ISBN 0-521-00531-0
- DAVIDSON, F., LYNCH, B.K. *Testcraft*. Yale University Press 2002. ISBN 0-300-09006-4
- DAVIES, A.; BROWN, A.; ELDER, C.; HILL, K.; LUMLEY, T.; McNAMARA, T. *Studies in Language Testing, Dictionary of Language Testing*. Cambridge University Press 1999. ISBN 0-521-65876-4
- Developing and teaching a course: Goals and Objectives*. [online] [cit.2006-08-11]. Dostupné na WWW: <http://www.pitt.edu/~ciddeweb/>
- DUBIN, F., OLSHTAIN, E. *Course Design*. 7th ed.: Cambridge University Press 1994. ISBN0-521-27642-X

- FCE (First certificate in English). [online] [cit.2006-08-11]. Dostupné na WWW:  
<http://www.cambridgeesol.org/exams/fce.htm>
- FULCHER, G. *Testing Second Language Speaking*. Pearson Longman 2003. ISBN 0-582-47270-9
- Goals and objectives*. [online] [cit.2006-08-11]. Dostupné na WWW:  
[http://www.lib.utexas.edu/services/instruction/tips/cp/cp\\_goals.html](http://www.lib.utexas.edu/services/instruction/tips/cp/cp_goals.html)
- HEATON, J.B. *Writing English Language Tests*. 4th impression: Longman Group 1991. ISBN 0-582-00237-0
- HUGHES, A. *Testing for Language Teachers*. 4th printing: Cambridge University Press 1992. ISBN 0-521-27260-2
- CHRÁSKA, M. *Didaktické testy*. Paido Brno 1999. ISBN 80-85931-68-0
- KET (Key English test). [online] [cit.2006-08-11]. Dostupné na WWW:  
<http://www.cambridgeesol.org/exams/ket.htm>
- Katalog požadavků ke společné části maturitní zkoušky v roce 2004, anglický jazyk, MŠMT 2000*
- Katalogy požadavků ke zkouškám, zadávaným Ministerstvem v rámci maturity*; [online] [cit.2006-08-11]. Dostupné na WWW: <http://www.ceremat.cz/katalogy/>
- LUOMA, S. *Assessing Speaking*. Cambridge University Press 2004. ISBN 0-521-80487-6
- LYNCH, T. *Study Listening*. 13th printing: Cambridge University Press, 1st published 1983. ISBN 0-521-27314-5
- MAGER, R.F. *Preparing Instructional Objectives*. 2<sup>nd</sup> ed.: Fearon Publishers, Inc., Belmont, California 1975. ISBN-0-8224-5601-X
- MC NAMARA, T. *Measuring Second Language Performance*. Longman 1996. ISBN 0-582-089077 PPR
- MUELLER, J. *Authentic Assessment Toolbox: Standards*. [online] [cit.2006-08-03]. Dostupné na WWW: <http://jonathan.mueller.faculty.noctrl.edu/toolbox/standards.htm>
- MUELLER, J. *Authentic Assessment Toolbox: Types of Standards*. [online] [cit.2006-08-07]. Dostupné na WWW: <http://jonathan.mueller.faculty.noctrl.edu/toolbox/standardtypes.htm>
- MUELLER, J. *Authentic Assessment Toolbox: What is Authentic Assessment?* [online] [cit.2006-08-07]. Dostupné na WWW: <http://jonathan.mueller.faculty.noctrl.edu/toolbox/whatisit.htm>
- NUNAN, D. *Syllabus Design*. 11<sup>th</sup> ed.: Oxford University Press 2002. ISBN 0-19-437139-5
- RICHARDS, J.C., PLATT, J. & PLATT, H. *Dictionary of Language Teaching & Applied Linguistics*. 8th edition Harlow, Longman Group UK Limited 1999. ISBN 0-582-07244-1

- SHOHAMY, E.: *The Power of Tests*. 1st ed. Longman 2001. ISBN 0-582-42335-X
- TRNA, J. Vzdělávací standardy pro základní a střední školy. *Pedagogika*, 1996, roč. XLVI, s. 349 – 353
- Standard vzdělávání ve čtyřletém gymnáziu*. Praha, Fortuna 1999. [online] [cit.2006-08-08]. Dostupné na WWW:  
[http://www.vuppraha.cz/download.php?f=Vestnik\\_MSMT\\_04\\_1996.pdf](http://www.vuppraha.cz/download.php?f=Vestnik_MSMT_04_1996.pdf)
- Standardy pro pedagogické a psychologické testování*. 1. české vydání. Testcentrum Praha 2001. ISBN 80-86471-07-1
- Improving America's School: A Newsletter on Issues in School Reform - Spring 1996* [online] [cit.2006-08-11]. Dostupné na WWW:  
<http://www.ed.gov/pubs/IASA/newsletters/standards/pt2.html>
- UNDERHILL, N. *Testing Spoken Language*. 8th printing: Cambridge University Press 1994. ISBN 0-521-31276-0
- User's Manual for the ITEMAN Conventional item Analysis Program*. Second Edition, Windows 3.x/95/NT Version. Assessment Systems Corporation 1989-1996
- VÁLKOVÁ, J. Český a americký způsob zpracování vzdělávacích cílů a obsahu vzdělávání, *Pedagogika* 1999, roč. XLIX, s. 236 – 248
- VANDERGRIFT, L. Listening: Theory and practice in modern foreign language competence. In *Guide to Good Practice for Learning and Teaching, Linguistics and Area Studies* [online] [cit.2005-11-02]. Dostupné na WWW:  
<http://www.lang.ltsn.ac.uk/resources/goodpractice.aspx?resourceid=67>
- [http://www.sbe.wa.gov/reports/reports/CAAFinalRPT/FNL%20CAA%20RPT.APPENDIX%20P%20AERA%20Pos.%20Statement .pdf](http://www.sbe.wa.gov/reports/reports/CAAFinalRPT/FNL%20CAA%20RPT.APPENDIX%20P%20AERA%20Pos.%20Statement.pdf) 1.8.2006
- <http://www.ed.gov/pubs/IASA/newsletters/standards/pt1.html> 8.2.2007

## **BIBLIOGRAFICKÝ ZÁZNAM**

SIKOLOVÁ, M. *Zjišťování výsledků vzdělávání v anglickém jazyce*. Praha, 2007. 165 s.  
Doktorská disertační práce. Univerzita Karlova v Praze. Pedagogická fakulta.

## **ANOTACE**

Práce se zabývá stanovováním výukových cílů a formulací standardů v předmětu anglický jazyk. Věnuje pozornost metodám testování čtyř základních řečových dovedností. Na základě kritického zhodnocení Katalogu požadavků ke společné části maturitní zkoušky v roce 2004 navrhuje alternativní katalog, jehož konkretizací jsou čtyři subtesty k ověření dovednosti poslechu, čtení, psaní a mluvení v rámci maturitní zkoušky. Dva ze subtestů jsou ověřeny v praxi a získané výsledky jsou komentovány a interpretovány.

## **KLÍČOVÁ SLOVA**

výukové cíle – vzdělávací standardy - testování jazyků – řečové dovednosti – testovací techniky – validita - reliabilita

## SUMMARY

As educational policy-makers in developed countries put effort to improve their educational systems, they often focus their attention on correct formulation of educational goals, objectives and standards. At the beginning of the first (theoretical) part, the dissertation mentions and comments on most respected views concerning this area. Essential ideas of Mager and Bloom are presented and discussed. Since the dissertation deals with the assessment of educational results, the importance of setting clear and understandable requirements is being stressed, as they are a pre-requisite for a valid and reliable assessment. The considerations about necessary features of a good test deal with validity, reliability, objectivity and authenticity.

Next, as the dissertation looks into assessing English as a foreign language, it briefly characterizes current internationally recognized standardized language tests (KET, PET, FCE, IELTS, CAE, CPE, TOEFL). Further on, it presents the experts' views on defining language skills and tests that measure them. Testing techniques for both receptive and productive skills are described, highlighting their main advantages and disadvantages. To conclude the first part, a table was designed in an attempt to classify language testing techniques according to cognitive process categories of Bloom's taxonomy. The intention of designing the table was to bridge the knowledge of teaching specialists and language testing specialists.

As far as the second part of the dissertation is concerned, its main objectives were to critically evaluate an educational document referring to requirements for the school-leaving exam in the Czech Republic published in 2000, to design an alternative document eliminating the shortcomings of the previous one, to design a sample test following the alternative model, to pre-test two of the subtests as well as to analyse the process and the results of piloting.

The evaluation of the document specifying requirements for school-leaving exam has shown both positive and negative features of it. A survey carried on six different high schools in Brno has revealed deep differences in administration as well as in assessment of English language in school-leaving exams. Following the findings of document evaluation and school survey, a conclusion has been made that there are many questions in testing English during the school-leaving exam to be addressed. That is why an alternative document specifying required knowledge and skills for two difficulty levels of school-leaving exam has been designed. This document has been evaluated by nine respondents (high-school teachers, parents, university teachers) in order to get at least a brief feedback. The document seems to be clear enough and the requirements are perceived as reasonable.

Bearing in mind the newly-designed document as well as basic principles of testing theory, four subtests have been designed to assess basic language macro-skills on pre-intermediate level roughly corresponding to A2 level of Common European Framework of Reference. All the four tests have been commented on by two testing specialists and their suggestions have been taken into consideration.



Two of the subtests, reading and writing, were administered to altogether 126 both high-school and university students during the period of time ranging from November 2005 to January 2006. The process of assessing both subtests was given a lot of attention. While the reading subtest, mostly scored objectively, was assessed exclusively by the author, the writing subtest was always scored independently by two assessors.

The results of the reading subtest have been analysed by computer software ITEMAN that has shown the basic characteristics of the test as a whole, as well as it has produced characteristics for each item: its difficulty, discrimination and distractors' functioning. Overall the test characteristics can be considered as satisfactory – the mean being 20.73 (maximum 30 points), the distribution being bimodal that enable dividing of the results as pass and fail, and reliability 8.444. Only two of the items are too easy (item difficulty 0.98 and 0.95) having consequently discrimination index too low (0.06). These items, if used in „live tests“, should be either discarded or re-written. In the writing subtests, the emphasis has been put on using the criteria by assessors. Most of the writing tasks were scored very similarly, with the correlation between the author and other assessors being 0.96. Feedback from the administration and assessment of both subtests has been collected and commented on.

As no research on lower scale can produce fully satisfactory results, other steps to be taken in research can be suggested. The document should be assessed and commented on by a larger number of high-school and university teachers to give a broader feedback. The most problematic issues could be then addressed in semi-structured interviews conducted with the teachers.

However, the tests themselves would require most attention – not only listening and speaking subtests that have not been pre-tested, but also the ones already piloted on a smaller sample of candidates – reading and writing subtests. The results of the listening subtest should be assessed by item analysis. Piloting the speaking test would require systematic observation and evaluation of both test tasks functioning, conducting the exam and assessing candidate's performance. Assessors' and interlocutors' training should be a must.