

## **Posudek oponenta disertační práce Ing. Evy Píšové**

### **“Pozitivní aspekty státní maturitní zkoušky z hlediska didaktiky AJ. Vliv testování psaní na výuku AJ na českých středních školách”**

**předkládané v roce 2018 na Ústavu anglického jazyka a didaktiky FF UK**

#### **I. Stručná charakteristika práce**

Posuzovaná disertační práce je zaměřena na hodnocení písemného projevu v anglickém jazyce a možnosti jeho rozvoje v rámci výuky angličtiny na středních školách. Na teoretickou reflexi tématu v první polovině práce navazuje část empirická, ve které autorka těží poznatky ze tří výzkumných postupů: 1) Analyzuje rozsáhlá data ze společné části maturitní zkoušky za několik po sobě jdoucích let jejího konání v jarním termínu. Tato data, resp. řadu dílčích výsledků, autorka získala od Centra pro zjišťování výsledků vzdělávání, které společnou část maturity realizuje. 2) Sama provedla didaktický experiment malého rozsahu. 3) Oslovila učitele angličtiny na středních školách s maturitním oborem a získala prostřednictvím dotazníku jejich názory vztahující se k výuce písemného projevu.

#### **II. Stručné celkové zhodnocení práce**

Autorka ve své práci těží ze spolupráce s Centrem pro zjišťování výsledků vzdělávání (CZVV) i ze svých zkušeností učitelky angličtiny na střední škole. Díky spolupráci s CZVV (podílí se na metodické přípravě učitelů i hodnocení maturitních písemných prací ve funkci vedoucí hodnocení předmětu) má mimořádný vhled do systému hodnocení písemných prací v anglickém jazyce u státní maturity i do jeho postupných proměn. Současně se jí podařilo získat data či analýzy, které nejsou veřejnosti přístupné a to ani veřejnosti odborné. Tyto analýzy přinášejí unikátní poznatky využitelné pro didaktiku anglického jazyka (pro výuku směřující k rozvoji dovednosti psaní, pro hodnocení písemného projevu). Didaktický experiment realizovaný samotnou autorkou práce nemůže díky svému rozsahu (22 žáků) přinést relevantní zobecnitelné výsledky, nicméně je metodologicky inspirativní a může být výzvou pro další badatele jej zorganizovat na větších souborech žáků a s vyvarováním nedostatků, které budou níže zmíněny. Pozice autorky ve spolupráci s CZVV umožňuje kontakty s řadou učitelů angličtiny ze středních škol s maturitou. Důvěryhodnost autorky se promítla do téměř 100% návratnosti dotazníků z oslovených učitelů. I tyto názory učitelů jsou důležité pro dokreslení tématu, neboť nám vypovídají o prostředí, v jakém se naši žáci učí angličtinu. Tyto znalosti mohou přispět k efektivněji zaměřené podpoře učitelů, jejich přípravě a dalšímu vzdělávání.

Na disertační práci zejména oceňuji její snahu o přesah do praxe. Autorka směřuje k didaktickým doporučením pro výuku angličtiny. Formuluje je v závěrech práce a při jejich formulaci těží z teoretické reflexe tématu, empirických výsledků i vlastních zkušeností.

#### **III. Podrobné zhodnocení práce a jejích jednotlivých aspektů**

Autorka správně uvádí, že „Přestože je vědecká a praktická závažnost vlivu zkoušek na výuku jazyka vysoká, je dosud u nás nedostatečně prozkoumána.“ (s. 9). Registrována je disertační práce Kolářové z roku 2016, která je zaměřena na problematiku výuky a hodnocení psaní v němčině na středních

školách. Škoda, že unikla pozornosti též disertační práce Novotné Knotkové z roku 2015, pracovnice CZVV, která je zaměřena též na němčinu, a má blízko k tématu posuzované disertační práce. Novotná Knotková se zabývala přímo písemným projevem u státní maturity. Konkrétně analyzovala reliabilitu hodnocení písemného projevu.

*K teoretické části práce mám dále jen několik méně podstatných poznámek a komentářů víceméně formulačního či formálního rázu.*

s. 12. Teze typu: „Přesto mnozí učitelé stále nemají dostatek zkušeností, metodických podkladů ani chuti psaní vyučovat.“ by měly být doloženy odkazem na informační zdroj. Obzvláště pokud jsou formulovány v rámci teoretických východisek práce.

s. 30. „Státní maturity se v České republice poprvé konaly ve školním roce 2010/2011, po čtrnácti letech příprav a mnoha diskuzích.“

„Státní“ byly maturitní zkoušky i před rokem 2011, byť se v mediálním a veřejném diskurzu používá tento přívlastek ve spojení s aktuálně uzákoněnou formou zkoušky, která má společnou část s jednotnou formou. V disertační práci a u tak silného tvrzení by bylo třeba se vyjadřovat přesněji.

s. 44. Grafy obr. 6 a 7 by bylo vhodnější udělat s relativními četnostmi, aby ty rozdíly byly mezi AJ a NJ srovnatelné. Vhodné by bylo držet i jednotnou formu grafů s těmi výše uvedenými.

s. 45. Důvodem nastolení jednotné úrovně u MZ nebylo jen snížení nákladů, jak se autorka domnívá. Kritizována byla a někdy stále je neujasněná koncepce (garance minimální úrovně vs. vypovídací hodnota pro VŠ či zaměstnavatele). V případě více úrovní zkoušky je to např. problém volby úrovně ze strany žáka na některých středních odborných školách a řešení případů nevykonání zkoušky vyšší úrovně.

V kapitolách 3.4.1 až 3.4.4 autorka seznamuje čtenáře s tím, jak je hodnocena písemná práce i protřednictvím vážených průměrů. I přes uvedené příklady mně není jasné, co tvoří váhy a jak jsou stanoveny.

s. 58. „Od roku 2016 tak hodnocení za tento oddíl lépe odpovídá skutečnosti.“ Co se myslí v této souvislosti „skutečností“?

s. 63. „Hodnotitelé netápou v oblasti gramatiky, mluvnické prostředky prezentované v různých učebnicích se totiž shodují, také splnění komunikačního cíle není bez použití vhodné mluvnice možné.“ K čemu se vztahuje „také“? Logická vazba není příliš jasná.

s. 83. Je třeba být přesnější ve formulacích, obzvláště u „silných“ tvrzení typu:

„Výstupní úroveň maturantů ze SOŠ se v souladu s RVP i podle výsledků dosažených v rámci SMZ165 nejvíce blíží zkoumané úrovni B1.“ Některých ano, některých ne, někteří jí dosahují jistě, někteří ji jistě nedosahují. Asi by byla vhodnější tvrzení typu: Většina žáků SOŠ požadované úrovně B1 dosahuje, ale vyskytují se i žáci, kteří této úrovně nedosahují. Tedy nehovořit o „úrovni maturantů“ obecně.

Důležitá tvrzení (poznámka pod čarou č. 166 „Reálná úroveň jazykových dovedností žáků pedagogického lycea zapojených do výzkumu odpovídá úrovni žáků SOŠ.“) je potřeba nějak doložit.

Stejně tak nestačí konstatování „Pro maximální srovnatelnost obou souborů byly voleny takové skupiny, které byly na stejné jazykové úrovni, měly stejnou vyučující a používaly stejné učebnice.“

*Připomínky k realizovaným výzkumům.*

1) Velikost empiricky zjištěného korelačního koeficientu nezávisí jenom na věcné souvislosti mezi měřenými konstrukty, ale i na variabilitě dat obou proměnných dané konstrukty reprezentující a na přesnosti jejich měření (reliabilitě). Z těchto důvodů není příliš snadné porovnávat korelační koeficienty mezi různými dvojicemi proměnných a vyvozovat z nich soudy o souvislostech konstruktů. Je tedy čistě metodologicko-statistickou zákonitostí, že test skládající se z řady vzájemně konsistentních položek bude vykazovat vyšší korelační koeficient s nějakou proměnnou, než izolované položky, nebo subtest s nízkým počtem položek. Podobně otevřené úlohy bývají zpravidla obtížnější než úlohy uzavřené a navíc absence, nebo alespoň výrazná eliminace možnosti tipování zvyšuje potenciál dobře formulované úlohy dosáhnout vyššího korelačního koeficientu s nějakou související proměnnou, než je tomu v případě položky uzavřené. Kdyby si autorka byla těchto zákonitostí vědoma, neformulovala by podle mého soudu minimálně sporné interpetace typu:

s. 98, bod 3. „Velmi překvapující je skutečnost, že nejvyšší míra korelace mezi hodnocením ve IVB a jednotlivými částmi didaktického testu byla zjištěna u 3. části DT (až 0,52 v roce 2016).

s. 101. „Velmi překvapující je skutečnost, že části DT zaměřené na poslech korelují s PP více než části zaměřené na čtení.“

s. 105. „Pokud získané hodnoty porovnáme s korelačními koeficienty vyjadřujícími souvislost písemné práce s celým subtestem čtení (graf 16), jsou korelační koeficienty získané pro úlohy zaměřené na postižení hlavní myšlenky dokonce výrazně nižší“

Proto jsou problematické i některé příliš jednoznačně formulované závěry:

s. 103. „hypotézu H je proto třeba zamítnout“.

Byť si tohoto problému autorka není plně vědoma, některé formulace nasvědčují tomu, že zde nějaký problém tuší, když diskutuje možný vliv přebodování některých úloh pro zpřesnění měřeného konstruktů a zastoupení otevřených úloh ve 3. části didaktického testu (např. na s. 98).

2) Experiment je dobře promyšlen z hlediska dvou vhodných skupin žáků, z hlediska použití dvou srovnatelných zadání písemné práce jako pretest a posttest a s jejich prohozením pro tento účel v obou skupinách. Počty žáků 10 plus 12 nedovolují z výsledků experimentu činit větší zobecňující závěry, nicméně způsob realizace experimentu může být inspirací pro podobně didakticky zaměřené experimenty na větších souborech žáků. Přesto bych navrhoval k realizaci a prezentaci takového experimentu následující úpravy:

1. Bylo by žádoucí, aby práce před i po hodnotil nezávislý hodnotitel. Nejlépe tak, aby nevěděl, které práce jsou pretestové a které posttestové. Případně aby ty práce hodnotili nezávisle dva hodnotitelé. Autorka v roli hodnotitelky může být podvědomě přísnější či mírnější v některých hraničních případech a neúmyslně, ale přesto ovlivnit výsledky experimentu podle svých přání. Dva či více hodnotitelů zase zpřesní měření daných konstruktů a umožní získat informaci o reliabilitě měření.

2. Ztěžujícím důvodem, proč byl experiment realizován, byla intervence mezi pretestem a posttestem. Detailní popis intervence však chybí. Právě to, co je z hlediska didaktiky i realizace experimentu klíčové, není čtenáři nabídnuto. Autorka se dopouští nepatřičných zobecnění (s. 114 „Linearita dějové linie je dána přirozeným dodržáním posloupnosti jednotlivých bodů zadání, na ostatní faktory působící

případnou nelinearitu textu nemá cíleně zaměřená příprava vliv.“) na adresu intervencí a čtenář neví, na základě jaké konkrétní intervence se tak činí.

3) Některé odpovědi v dotazníku mohly být ovlivněny „očekáváním žádoucí odpovědi“, např. „Výuce psaní dávám přednost před výukou ostatních dovedností.“, „Učím studenty v textu vhodně organizovat myšlenky“. Rozdíl je v tom, zda se ptáme na názor a výsledek správně jako názor interpretujeme, nebo se ptáme na skutečnost a tu se pak někdy mylně snažíme jako takovou interpretovat. Podle mého soudu je otázka na překážky při výuce psaní komplikovanější pro respondenty zadaná, než předchozí otázka.

10. Učím studenty pracovat s tištěnými slovníky.

rozhodně ano                      spíše ano                      nevím                      spíše ne                      rozhodně ne

**E. Na bodové škále 1 – 5 označte, do jaké míry pocítujete jako překážku při výuce psaní následující skutečnosti:** (1 – vůbec mi nevadí, 5 velmi mi vadí)

- |  |   |   |   |   |   |
|--|---|---|---|---|---|
| • časová náročnost při psaní v hodině  | 1 | 2 | 3 | 4 | 5 |
| • nedostatek materiálů pro výuku psaní | 1 | 2 | 3 | 4 | 5 |

Resp. nevím, jak vypadala formulace v elektronické verzi dotazníku, kde mohl být problém odstraněn, nebo naopak posílen. Škála je otočená – souhlasná odpověď je vpravo a popis škály je nad položkami. Je možné, že někteří respondenti si toho při rychlém čtení nevšimli, nebo to dále „zapomněli“. I tím by se možná dala vysvětlit některá tvrzení, např.:

s. 128: „Zajímavý fakt, že tytéž faktory považují někteří učitelé za obtěžující, zatímco jiným nevadí, dokumentují grafy 38 a 39.“

U H12 by bylo vhodnější statistické zpracování buď přes průměry škál, nebo kumulaci procent. Snáže by se daly výsledky porovnávat.

#### *Formulační nepřesnosti v empirické části*

s. 77. „Výzkumným problémem je kompetence psaní v anglickém jazyce ověřovaná písemnou prací v rámci SMZ z AJ.“ Podle metodologických publikací (Pelikán 2011, Chráska 2016, které jsou uvedeny v seznamu literatury) má výzkumný problém vyjadřovat vztah mezi dvěma či více proměnnými.

Výzkumné otázky jsou formulovány dobře. Jedinou výjimkou je VO10 („Jaká je vstupní úroveň žáků nastupujících na střední školy?“), která navozuje představu, že je tato vstupní úroveň zjišťována. Zjišťováno je pouze přesvědčení učitelů o tom, jakou úroveň žáci mají.

s. 80. „Toto šetření bylo doplněno vlastním empirickým sledováním pedagogické skutečnosti.“ Proč? Jak se vztahuje k formulovaným výzkumným otázkám? Jak bylo reálně uplatněno ve výzkumu a při interpretaci dat?

s. 89. „Velký význam pro zaručení objektivitu má i množství zpracovávaných dat.“ Velké množství dat není zárukou objektivitu, pouze zvyšuje možnosti zobecnitelnosti (externí validity).

s. 93. „Nulová hypotéza tak nebyla potvrzena, byla zjištěna závislost mezi výsledky didaktického testu a písemných prací v oblasti úspěšných PP, tedy těch, které dosahují 16 a více bodů.“ Nulová hypotéza

se nepotvrzuje. Nulovou hypotézu se buď podaří vyvrátit, jako v tomto případě, nebo se ji vyvrátit nepodaří, ale tím se nebere jako potvrzená. V případě VO1 by bylo vhodnější využít přímo korelačního koeficientu podobně jako u VO2. Je navíc níže spočítán a komentován. Bylo by to vhodnější, než prezentace grafem 7 na s. 92 (Co znamená v tomto grafu PP vlevo od 0 na vodorovné ose?).

s. 98. Prezentace grafy 9 a dále je v daném případě vhodnější, než prezentace dat stejného charakteru v grafu 8, kde navíc spojnice mezi body postrádají smysl.

s. 104. Není jasné, jak byly sloučeny dvě úlohy na postižení hlavní myšlenky. Byl vzat prostý součet bodů z těchto dvou úloh?

Některé poznámky, např. 175 na s. 106 jsou natolik metodologicky důležité (upřesnění analyzovaných dat a role autorky v procesu hodnocení), že mají být součástí hlavního textu.

s. 117. Je nejasné, jak byli rozděleni učitelé SOŠ a G pro analýzy, když je jich několik vyučujících na obou typech SŠ (Byli z analýz vyloučeni?, Kam byli případně přiřazeni? K SOŠ i ti co učí na SOU?).

s. 132. „...týmů odpovídá běžně používané praxi“. Co se rozumí „běžnou praxí“? V kontextu ČR se jedná o mimořádně nadstandardní praxi.

s. 135. „VO10: V rámci průzkumu postojů učitelů k výuce psaní, především v souvislosti se SMZ, byla zjišťována skutečná vstupní úroveň žáků“ „Skutečná“ ne, dále autorka interpretuje správně jako názory učitelů na úroveň žáků.

I přes uvedené formulační nepřesnosti chci vyzdvihnout, že shrnutí výsledků dotazníkového šetření v kapitole 7.3.7 je zdařilé. Formulace jsou bezchybné a adekvátní získaným datům.

#### *Obecné poznámky diskusní povahy.*

Autorka se oprávněně kriticky vyjadřuje k tomu, že „SMZ ...poskytuje každoročně i rozsáhlé soubory výsledků, kterých žáci v maturitních testech dosahují, se kterými se však ve větší míře v pedagogickém výzkumu dále nepracuje“. Problém však není na straně pedagogického výzkumu, ale CZVV, které neposkytuje data v takové podobě, aby se s nimi dalo dobře výzkumně pracovat. Pro CZVV nemůže být alibi, že „neuvolňuje citlivá data pro další zpracování“, jak autorka uvádí v poznámce pod čarou 171 na s. 93. Pedagogický výzkum by mohl těžit z dat, která by nebyla citlivá, jedinci by nebyli identifikovatelní, ale data by byla vzájemně propojitelná skrze anonymní kódy (z různých zkoušek, z různých částí zkoušek). Taková data však CZVV nenabízí.

Empirická část obsahuje tři oddělené výzkumné postupy (analýza dat z CZVV, vlastní experiment, dotazování učitelů). Design výzkumu i výsledky jsou i v dílčích podkapitolách strukturovány podle těchto postupů. Čtenář musí myšlenkově stále přeskakovat. Otázka je, zda by nebylo vhodnější držet popis výzkumu až po výsledky v rámci jednoho postupu a teprve pak provést společné shrnutí výsledků s praktickými doporučeními pro výuku.

Za velmi cennou pokládám kapitolu 7.1.4 v zaměření na neúspěšné práce. Z této analýzy vyplývají i praktická doporučení pro výuku.

Za velmi cenná pokládám doporučení pro praxi na s. 140 a dále. Škoda, že jsou tato doporučení „utopena“ v textu. Zasloužila by si nějaké strukturování a vytažení buď formou podkapitol nebo zvýrazněním písma klíčových slov.

#### **IV. Dotazy k obhajobě**

Jak je určováno hodnocení písemné práce s využitím vah za deskriptory? Jakým způsobem tyto váhy do hodnocení vstupují?

Jak byli rozděleni respondenti do skupin podle typu školy v analýzách výsledků dotazníkového šetření?

Na s. 142 je uvedeno: „Nečitelný rukopis se tak může v budoucnu stát problémem, který naruší srozumitelnost textu, a tím i sdělení komunikačního cíle.“ Nemělo by se tedy v budoucnu realizovat psaní písemné práce na počítači či tabletu? Jaké přednosti a hrozby vidíte v této možnosti?

#### **V. Závěr**

Řada kritických poznámek na adresu formulačních nepřesností v práci, či k některým aspektům výzkumu nemůže zastřít celkový přínos práce, který byl shrnut v celkovém zhodnocení práce na úvod posudku. Práce je po formální stránce zpracována přehledně včetně seznamu tabulek a grafů. Autorka vhodně a v hojné míře užívá zkratk, jejichž seznam je též uveden na začátku práce. Přílohy jsou voleny účelně.

Věřím, že přínos pro rozvoj didaktiky anglického jazyka bude s odstupem času ještě více oceněn. Kéž se podaří autorce a jejím následovníkům dále dobře vytěžovat data ze státní maturity a kéž se najdou podobně zapálení učitelé-výzkumníci, kteří budou mít odvahu se pouštět do promyšlených didaktických experimentů.

Předložená disertační práce splňuje požadavky kladené na disertační práci, a proto ji doporučuji k obhajobě a předběžně ji klasifikuji jako *prospěla*.

25.7.2018

PhDr. Martin Chvál, Ph.D.