

Oponentský posudek na disertační práci

Pavλίna Synková: Popis staročeské apelativní deklinace (se zřetelem k automatické morfologické analýze textů Staročeské textové banky)

Oponent: doc. Mgr. Miroslav Vepřek, Ph.D.

Posuzovaná disertační práce se vztahuje ke dvěma základním okruhům výzkumu – první můžeme označit jako paleobohemistický, tedy zaměřující se na analýzu morfologického subsystému substantivní deklinace ve staré češtině, druhý potom k oboru počítačové a korpusové lingvistiky. Ve svém posudku budu akcentovat především okruh první. Přesto na úvod uvedu dvě důležité připomínky i k okruhu druhému.

První z nich spočívá v důležitém upozornění na formální podobu analyzovaného materiálu. Nejde ani tak o složení a charakteristiku digitálního zdrojového korpusu jako celku, ale o jeho grafickou podobu, která je založena na **transkripci**. Autorka si problematičnost transkripce, jež je nutně zároveň interpretací editora daného textu, pro lingvistické zkoumání uvědomuje, vyjadřuje se takto např. na s. 33, nicméně přesto se domnívám, že relevantní pohled na historická jazyková fakta je plně možný pouze s veškerým respektem k jejich skutečnému doloženému zápisu. Za všechny případy problematických jevů můžeme označit neoznačování kvantity písaři užívanými staročeské varianty spřežkového (nebo samozřejmě i tím spíše primitivního) pravopisu. Jak je potom možné relevantně posoudit a počítačově zpracovat např. otázku krácení koncovek, srov. kupříkladu koncovku dativu plurálu o-kmenů (-om x -óm)? Je zřejmé, že tento mnou uvedený příklad (a autorkou probíraný např. na s. 57) je záležitostí spíše marginální, neboť obě koncovky se v historických mluvnicích popisují jako rovnocenné. Avšak v případě, že počítačové zpracování jazykových údajů má sloužit především jako nástroj pro jejich exaktní sumarizaci, dle mého názoru je jeden z těchto základních přínosů strojové lingvistiky přinejmenším problematizován. Přímo vzorovým případem je pak oddíl 3. 1 s názvem „Apelativa s alternací délky“. I když v úvodu této kapitoly autorka uvádí, že „značení kvantity nebylo ve sledovaném období nijak systematické a jednotlivé památky se liší v tom, zda vůbec, případně do jaké míry kvantitu značí,“ nejedná se o pouhou relativizaci, nýbrž o celkové zpochybnění smyslu takovéto analýzy na základě transkripce. Byť bude editor dané staročeské památky či redaktor digitalizované edice sebelepší a sebepečlivější znalec staré češtiny, pořád se bude jednat v případě značení kvantity o interpretaci, nikoli spolehlivě reálný a zapsaný jazykový stav. V neposlední řadě skutečnost, že za základ zkoumání je vzata transkribovaná podoba staročeských textů, vede autorku někdy až k téměř absurdním tvrzením, jako např. na s. 56 v poznámce č. 39, kde se uvádí, že je „ojediněle doložena i transkripce -uóv, doklad zařazuji do seznamu výjimek.“ K tomu poznamenávám, že formulace „doklad“ a „doložení“ užíváme pro podoby, které se reálně nacházejí ve zkoumaných textech.

Druhá připomínka vychází z mého poněkud skeptického pohledu, a je tedy do značné míry subjektivní. Přesto se ale nemohu ubránit přesvědčení, že soubor ručně tagovaných výjimek pro jednotlivá lemmata či morfémy je někdy tak velký, že pro diachronní zkoumání se mnohem efektivnější stále jeví léty prověřená „ruční“ metoda excerptce jazykového materiálu. Znovu však zdůrazňuji, že v tomto případě se může jednat o můj osobní postoj.

Nyní se již přesunu k druhému avizovanému základnímu okruhu, a to k problematice paleobohemistické. Na prvním místě je třeba konstatovat, že i přes svou rozsáhlou analytičnost a detailnost uvedených rozborů na velmi obsáhlém prostoru (disertace disponuje více než pěti sty stranami) postrádá dle mého názoru práce přesně definované cíle a především závěry. Ty jsou nahrazeny obsáhlým „shrnutím“, které ale pouze zhuštěně opakuje výčty uvedené v textu práce. Dále musím uvést, že přínos k poznání staročeské substantivní deklinace a vlastní vědecký výzkum v této oblasti, který má disertační práce ze své podstaty obsahovat, zde bohužel neshledávám prakticky žádný. Rozsáhlé pasáže věnované charakteristice deklinačních typů a podtypů nejsou ničím jiným, než kompilací poznatků získaných excerpcí odborné literatury (pro objektivitu dodejme, že zhusta doplněné o nové doklady z památek obsažených v elektronických zdrojích). Jisté dílčí vlastní poznatky autorky můžeme spatřovat snad ve vymezení některých deklinačních podtypů na základě zjištěných alternací.

Bohužel však i v tomto převážně kompilačním popisu staročeské deklinace nacházíme některé nepřesnosti. Zarážející jsou zejména nedůslednosti či nesprávné údaje v prezentovaných tabulkách, které mají představovat nejdůležitější výsledky práce doktorandky. Jmenujme např. chybějící (a vlastně základní!) staročeskou nulovou koncovku nominativu singuláru ženských i-kmenů v tabulce č. N1a, chybějící koncovku nominativu sg. њja-kmenů v tabulce M1a vytvářející dojem, jakoby tento základní tvar vůbec nebyl doložen (!), dále v tabulce J u shrnutí koncovek t-kmenových substantiv charakteristiku variantní koncovky dativu singuláru *-u* jako u-kmenové aj. Celkově tyto tabulky, koncentrující synchronní a diachronní pohled na dubletní koncovky staročeské deklinace, lze jistě považovat za dobrou ideu, které může v konečném důsledku zprostředkovat „mechanizovaný“ pohled a podklad pro strojové zpracování, avšak bylo by žádoucí zpracovat tabulky s větší pečlivostí a důsledností. V tomto kontextu uvedme ještě např. tabulku A1a, kde nulová koncovka v nominativu a akuzativu singuláru může být stejně tak původně o-kmenová jako u-kmenová (tato druhá skutečnost není reflektována); dále v téže tabulce není úplný výčet možných koncovek nominativu plurálu, což se týká i následující tabulky A1b přinejmenším u „jmen osob“, podvzor „pán“.

Nepřesnosti a nedůslednosti se bohužel objevují i ve shrnujících tabulkách šesté kapitoly. Jako příklad vybírám tabulky v oddílu 6.4.2 s názvem „Přehled hláskových změn“. Zde hned první údaj první tabulky (č. 32) charakterizuje změnu *ú > ou* jako vyskytující se pouze na začátku slova (užití výrazu „slovo“ zde považuji za terminologicky nepřesné – mělo by být přinejmenším uvedeno, ve smyslu které z definicí je zde míněno) s příklady *ústa/ousta*. Jak do této charakteristiky zapadá namátkou dvojice *múka/mouka*, *súd/soud* a mnoho jiných? Navíc by dle logiky popisků tabulky měla být tato změna údajně „pouze v tvarotvorném základu“, avšak nalezneme ji také níže v tabulce č. 34, jež obsahuje „změny, jež se vyskytují v tvarotvorném základu i v koncovkách. Ještě problematičtěji hodnotí zařazení změny „ě = e, ie = é po l“ do tabulky č. 33. V první řadě se domnívám, že se tím odporuje vytyčenému cíli, že „popis změn počítá s podobami koncovek k roku 1300“ (s. 495, bod 5),¹ neboť ztráta jotace po *l* je datována před tento stanovený časový bod. Dále je však třeba poznamenat, že tato tabulka zahrnuje „hláskové změny, jež se vyskytují pouze v koncovkách“. Znamená to tedy, že např. ve slovech *mléko* (< **mlěko*), *léto* (< **lěto*) a mnohých jiných se realizovala daná hlásková změna v koncovce?

¹ Ostatně i toto časové vztážením považuju za velmi odvážný cíl, bez dalšího vymezení přinejmenším problematický. Zejména se nabízí otázka, na základě čeho je zjištěn stav k danému roku, zda dle datece historických mluvnic, nebo jazykové podoby památek zachycených písmem s více či méně rozkolísanou „normou“.

Při ústní obhajobě práce bych považoval za důležité, aby doktorandka zodpověděla taktéž některé mé otázky, které vycházejí z jejích vlastních formulací:

1. Co vás vede k poněkud neobvyklému označení „malé a velké deklinační typy“ na s. 48? Jedná se o vlastní „termín“, či vycházíte ze sekundární literatury (pak by bylo žádoucí doplnit odkaz)?
2. V jakém smyslu považujete i-kmeny za „centrální deklinační typ apelativ ženského rodu“? Je to na základě frekvence, pozice v systému, produktivity aj.?
3. Za důležité považuji též, aby bylo specifikováno a odůvodněno tvrzení v samotném úvodu práce na s. 23, že „morfologický systém staré češtiny se od systému češtiny současné zásadním způsobem liší.“

Formální podobu práce považuji za zdařilou, což se týká nejen grafického provedení, ale též stránky stylistické. Drobné nedůslednosti jsou spíše ojedinělé, jako např. nezařazení hojně užívané zkratky NT do seznamu zkratek a značek, v rámci kterého mi není zřejmě vřazení zkratek „počátek“, „polovina“, „staročeský“, „století“ mezi „zkratky pro zdroje odborné literatury a textové banky“.

Z výše uvedeného dle mého názoru vyplývá, že celkové hodnocení předložené disertační práce Pavlíny Synkové považuji za problematické. Nelze zpochybnit její jisté kvality, a to zejména značně novátorský přístup k automatickému zpracování jazykového materiálu staré češtiny. Na druhou stranu však vyjadřuji velké pochybnosti nad zvolenou základní podobou analyzovaných lingvistických údajů, tedy formou transkribovanou. Tato skutečnost dle mého názoru může přinášet někdy až významně zkreslené výsledky zkoumání (viz např. výše zmíněný případ s kvantitou vokálů). Diachronní lingvista se musí ve svých analýzách spolehnout především na skutečně doložené tvary v celé jejich potencionální rozkolísanosti a nedokonalosti pramenící z různých grafických systémů a jejich specifik. A konečně, klade-li si disertace za cíl podat podklady pro exaktní strojové zpracování, měla by sama být zhotovena s přesností a důsledností.

Jen s velkými výhradami tedy práci **doporučuji** k obhajobě a předběžně klasifikuji jako **prospěla**. Pro ústní řízení při obhajobě práce zejména doporučuji, aby doktorandka přednesla výstižné zdůvodnění přínosu a významu výsledků své disertace, a to nejen pro počítačové zpracování vybraného korpusu staročeských transkribovaných textů, ale též pro zkoumání staré češtiny v obecnější rovině.

V Olomouci dne 13. dubna 2017

doc. Mgr. Miroslav Vepřek, Ph.D.