

Komparativ v korpusu: explanace morfemické struktury českého stupňování na základě frekvence tvarů*

JAN KŘIVAN

Comparatives in the corpus: An explanation of the morphological structure of comparison in Czech on the basis of frequency of forms

Abstract: The aim of this paper is to thoroughly investigate the morphological marking of comparative and superlative synthetic forms of Czech adjectives and to show the role of frequency in explaining their structure. Based on functional usage-based approaches (the role of frequency effects and the relationship between frequency asymmetries and universal morphosyntactic asymmetries) and on natural written data from the Czech National Corpus, eight hypotheses on the relationship between frequency and the formal marking of Czech comparison forms were formulated. The main part of the text consists of a detailed analysis and examination of the given hypotheses, and, additionally, a complete classification of comparative forms in order to explore some of the frequency effects. The paper reaches two major conclusions: 1) the reducing frequency effect plays a significant role among comparative forms, and 2) frequency contributes to an explanation of the formal relations in individual types (classes of comparative forms). In general, frequency is revealed to be a good analytical tool which may, under controlled conditions, indicate and explain why language structures are the way they are.

Key words: comparison of adjectives, frequency, comparative, functional explanation, frequency effects, language corpus

Klíčová slova: stupňování adjektiv, frekvence, komparativ, funkční explanace, frekvenční efekty, jazykový korpus

V tomto příspěvku se budu zabývat stupňováním českých adjektiv, a to z hlediska různých způsobů jeho morfologického vyjádření ve vztahu k četnosti výskytu stupňovaných tvarů. Kladu si přitom cíle jak deskriptivní, tak explanatorní. Pokud jde o vlastní popis, představím zde celistvou strukturní a frekvenční analýzu stupňovaných syntetických tvarů adjektiv založenou na datech z Českého národního korpusu. Z teoretického hlediska se pak budu snažit pomocí tohoto průzkumu ověřit, zda a do jaké míry lze prostřednictvím frekvence vysvětlovat a predikovat morfemickou strukturu stupňování adjektiv. Vycházet přitom budu z poznatků funkčně zaměřených přístupů, především ze studií Bybeeové (2007b) a Haspelmatha (2008a, 2008b), jejichž hlavní teze v úvodu představím.

Stupňování českých adjektiv patří mezi dobře probádané oblasti českého jazyka, dosud nejvíce zeširoka se mu z pozic českého strukturalismu věnoval ve své monografii Marvan (2008), jehož cílem bylo postihnout také diachronní a evropský rozměr

* Za podnětné připomínky k předešlým verzím tohoto textu patří můj dík Viktoru Elšíkovi, Janu Chromému, Janu Křivanovi st., Evě Lehečkové a oběma anonymním recenzentům Slova a slovesnosti. Vznik článku byl podpořen z projektu MSM-0021620825 Jazyk jako lidská činnost, její produkt a faktor.

českého stupňování.¹ Základní popis, obvykle spojený s výčtem nepravidelných tvarů komparativu, je samozřejmou součástí každé české mluvnice. Přitom však mezi lingvisty neexistuje shoda, zda stupňované řady tvarů řadit do slovtvorby, nebo do morfologie.² V této souvislosti připomínám důkladnou formální a sémantickou analýzu českého stupňování od Karlíka a Hladké (2004), v níž autoři nacházejí dobré důvody pro jedno i pro druhé řešení.

Zaměření tohoto článku je odlišné. Chtěl bych v něm ukázat, jak stupňování, ale i jiné gramatické otázky mohou být nahlíženy, popisovány a vysvětlovány, podívám-li se na ně z jiné (v českém prostředí ne příliš rozšířené), funkčně-typologické perspektivy. V rámci toho představím také data, která mohou přispět k přehodnocení noremnosti některých stupňovaných tvarů. V neposlední řadě tímto příspěvkem naznačím další možnosti a způsoby, jak lze využívat jazykové korpusey k lingvistické práci.

Struktura článku je následující: v 1. oddílu představím hlavní teoretická východiska (především otázku tzv. frekvenčních efektů), na nichž je založen tento článek, v 2. oddílu podrobně popíšu metodiku práce s korpusem při přípravě a zpracovávání dat, v 3. oddílu formuluji hypotézy nad frekvenčními efekty a představím základní strukturní a frekvenční charakteristiky českého stupňování a v 4. oddílu ověřím hypotézy na materiálu českého jazyka, součástí tohoto oddílu je i přehled všech nepravidelných stupňovaných tvarů doložených v korpusech, v 5. oddílu článek uzavřu shrnutím nejdůležitějších závěrů.

1. Teoretická východiska

Studium frekvence a především její role v gramatice úzce souvisí s rozvojem funkčně zaměřených přístupů v americké a západoevropské lingvistice. V této větvi lingvistiky se od 60. let buduje funkcionalismus, který se od pražského funkčního strukturalismu odlišuje jak v postulovaných teoretických východiscích, tak zejména v rozdílném způsobu uchopení dynamičnosti jazyka (v synchronním i diachronním smyslu). Lingvisté pohybující se v tomto teoretickém rámci, který je postaven do značné míry na poznacích jazykového typologického srovnávání, se jednak věnují gramatice ve vzájemném propojení s jazykovým užíváním (popisují a vysvětlují, jak se gramatické struktury utvářejí v rámci komunikačního procesu), jednak se zabývají studiem gramatikalizace, tedy gramatiky v jejím dynamickém historickém vývoji. Z jejich pohledu se předpokládá, že gramatika vzniká historickým vývojem tím, jak se určité diskurzvní vzorce (které zohledňují potřebu komunikace) obecně a často používají. Zjednodušeně řečeno, „gramatiku vytváří konvencionalizace struktur prostřednictvím opakování“ (Bybee, 2007b, s. 6, vlastní překlad).³

¹ Marvan přímo hovoří o „evropské univerzálii“, slovo „evropské“ je na místě, neboť stupňování adjektiv pomocí afixů patří mimo evropské jazyky k velmi vzácným jevům (viz např. Cuzzolin – Lehmann, 2004, s. 1215).

² Naposledy Cvrček a kol. (2010, s. 126) zařazují ve své koncepci mluvnice stupňování mezi gramatické kategorie.

³ Z českých lingvistů se nezávisle na těchto funkčních přístupech soustavně zabývá metodologií procesuálně-komunikačního přístupu k řeči a jazyku J. Kořenský (viz např. 1987, 1994; k širšímu mezinárodnímu

Od konce 70. let bylo toto základní teoretické pole funkčních přístupů dále rozšiřováno (jak z hlediska variability dokladů, tak i horizontů lingvistického zkoumání), gramatika se začala zkoumat v rámci kognice a v souvislosti s tím se začaly prověřovat i možné efekty opakování (frekvence) na kognitivní reprezentaci jazyka (Bybee, 1985) a frekvencí začaly být predikovány typologicky univerzální vlastnosti jazykových struktur (Haspelmath, 2008b; viz také Croft, 1990). Tito lingvisté tak v současnosti rozvíjejí frekvenční studium Zipfovo (1949) a navazují i na kvantitativní lingvistiku 60. let a Greenbergovy (1966) typologické generalizace poznatků o frekvenci a jejím vztahu k příznakovosti.

Role zkušenosti a frekvence se promítla do konstrukce tzv. modelů založených na užívání (*usage-based models*), viz např. Bybeeová (1985). V těchto modelech jazyka jsou lexikon a gramatika vzájemně promísены – na struktury obou se nahlíží jako na velmi podobné a přístupné. V pozdějších studiích je také jazyk, jakožto jedna z kognitivních domén, pojímán jako tzv. vznikající systém (*emergent system*), v němž opakování je jeho nezbytným komponentem během vytváření organizované struktury. Gramatika bývá obrazně přirovnávána k pískové duně (která tvoří celek, ale je v neustálém procesu vzníkání a měnění a není ukotvena v čase ani prostoru), viz např. Hopper (1987) a jeho pojetí tzv. *emergent grammar*. Změny v gramatice a lexikonu pak mohou být v rámci takto rozšířeného kognitivního horizontu pojímány jako frekvenční efekty prostřednictvím kognice, tedy jako kognitivní odpovědi na opakování.

Výhodou uvedených přístupů pro lingvistický popis je skutečnost, že jsou jejich prostřednictvím systematicky analyzovatelné, vysvětlitelné a predikovatelné jevy jako supletivismus a další nepravidelnosti, které jsou ve formálně zaměřených teoriích nebo při tradičním způsobu popisu (např. v gramatikách) obtížně uchopitelné.⁴ Jak se ukáže dále, pro jejich použití není přítom bezpodmínečně nutné přistupovat na některá kognitivistickými předpoklady zatížená stanoviska.

1.1. Frekvenční efekty utvářející gramatiku

Působením frekvence na jazyk se dlouhodobě zabývá především americká lingvistka Joan L. Bybeeová. Budu vycházet z jejího úvodu (2007b, s. 9) k tematickému souboru starších i novějších příspěvků (2007a), v němž představuje několik frekvenčních efektů. K jejich formulaci Bybeeová dospěla v průběhu let na základě rozličného studia frek-

kontextu viz zejm. Kořenský, 2003), dále viz také Čech (2005, 2007). – Velmi kritické srovnání celé skupiny funkčně-typologických přístupů ve vztahu k českému strukturalismu podává Vykypěl (2009): Vykypěl tvrdí, že jejich některá východiska i mnohé myšlenky považované za inovativní jsou součástí již funkčního strukturalismu; např. (na s. 56) uvádí, že tři vlastnosti funkčně-typologického přístupu deklarované W. Croftem (komparativnost, funkčnost, historičnost) jsou aplikovatelné i na pražský funkční strukturalismus. Vykypělovu kritiku a jeho teze o znovuobjevování myšlenek Pražské školy považují jen částečně za oprávněné, věrohodná polemika však překračuje možnosti tohoto článku.

⁴ Jinými slovy lze říci, že v rámci těchto funkčních přístupů se vybudovaly teorie, které operují s frekvencí jako prostředkem k analýze, rozvedení a vysvětlení těch principů v jazyce, které se v lingvistických popisech vyskytují v pojmech prototypová kategorie (např. Rosch, 1977; Taylor, 1989) nebo centrum a periferie (Daneš, 1966).

vence v různých jazycích. Problematiku frekvenčního působení autorka osvětluje za terminologického rozlišení mezi tzv. frekvencí tokenovou a typovou.⁵ Pojmy budu ilustrovat i příklady z vlastního výzkumu českého stupňování.

T o k e n o v á f r e k v e n c e udává počet výskytů určité jednotky v textu. Jednotkou může být např. konkrétní hláska, slabika, slovo, fráze i celá věta. U stupňování lze např. zjišťovat „počet všech výskytů jednotky komparativu *větší*“. Vzhledem k tomu, že ve výzkumu zaměřeném na stupňování nemusím rozlišovat flektivní morfologické kategorie, definuji jednotku jako pozitiv/komparativ/superlativ ve všech jeho pádech, rodech a obou číslech, v dalších částech článku budu ekvivalentně používat termín „tvar“ nebo „pozitiv“/„komparativ“/„superlativ“, pokud je to z kontextu zřejmé.

T y p o v á f r e k v e n c e se oproti tokenové frekvenci týká jazykových struktur. Udává, v kolika rozdílných jednotkách je přítomna zkoumaná struktura. Při výzkumu českého stupňování bude např. zajímavé sledovat, v kolika různých adjektivních lemmatech se v češtině komparativ tvoří pomocí sufixu *-ejší*. Termín „lemma“ v celém článku používám ve významu „soubor tvarů komparativu, pozitivu a superlativu od jednoho adjektiva“.

Bybeeová (2007b) identifikuje tři efekty frekvence tokenové (i–iii) a jeden efekt frekvence typové (iv):

(i) **K o n z e r v a č n í e f e k t** spočívá v tom, že nepravidelné struktury s vysokou frekvencí mají tendenci uchovávat si svoji nepravidelnost, a nepodléhat tak vlivu produktivních vzorců. Např. v angličtině vysoce frekventovaná slovesa dosud tvoří nepravidelný minulý čas (např. *kept* ‚držel‘), zatímco málo frekventovaná slovesa ze stejné třídy již podlehla regularizaci (např. *wept* ‚plakal‘). Kognitivně tento efekt Bybeeová vysvětluje tak, že opakováním se posiluje reprezentace v paměti, vysoce frekvenční jednotky se tak automatizují (uživatelé jazyka je mají tzv. pevně zakořeněné, *entrenched*), a proto nepodléhají změnám. U českého stupňování se nabízí prozkoumat hypotézu, zda v rámci kategorie lemmat se stejným kmenovým zakončením zůstávají ta s vyšší frekvencí komparativu a/nebo superlativu nepravidelná, zatímco ta s nízkou frekvencí se zpravidelňují.

(ii) **R e d u k č n í e f e k t** je založen na pozorování, že vysoce frekventované jednotky mají tendenci podstupovat fonetickou redukci (míněna je jak redukce výslovnosti na úrovni hlásek, tak fonologická redukce hláskového inventáře jednotek a jeho zjednodušování) častěji a radikálněji než jednotky s nižší frekvencí. Působením tohoto efektu mohla vzniknout řada českých komparativů, např. se předpokládá, že místo staročeského tvaru *starější* se prosadila redukováná podoba *starší* (viz např. Gebauer, 1895, s. 60; Marvan, 2008, s. 175), zatímco u méně častých tvarů se stejným zakončením se neprosadila nebo vůbec nevytvořila.⁶

⁵ Při překladu pojmů do češtiny jsem se rozhodl zachovat terminologii, která je založena na užitečných termínech *token* a *type*.

⁶ Bybeeová fonetickou redukci vysvětluje v souvislosti s neuromotorickým zpracováním: „...opakované sledy neuromotorických příkazů a činností mají tendenci k tomu, aby byly zpracovány jako samostatné jednotky, a zároveň mají sklon k větší efektivitě, a to tak, že zapojené pohyby se více překrývají a redukují“

Ačkoliv se může zdát, že redukční a konzervační efekty působí proti sobě, podle Bybeeové (2007b, s. 13) je to jen zdánlivé, protože každý z efektů se týká různého typu jazykové změny.

(iii) *E f e k t a u t o n o m i e* je možné pojímat jako extrémní případ konzervačního efektu, Bybeeová však má důvod tento jev vydělit zvlášť. Autonomie (slov, případně frázi) je pojímána jako míra pravděpodobnosti, že slovo (tvar) bude reprezentováno ve slovníku mluvčího jako celá a separátní jednotka. Vedle frekvence autonomie ovlivňuje i sémantická jednoduchost a morfonologická pravidelnost (autonomnější jsou jednotky složitější a nepravidelné). Předpokládá se, že vysoce autonomní slova mají kognitivně slabší propojení s ostatními vztaženými slovy (jak v rámci paradigmatu, tak v rámci lexikální třídy). Podle tohoto pojetí tak mohou být i některé tvary v rámci morfologického paradigmatu autonomní a součástí lexikonu.

Nejzajímavějším důsledkem autonomie je suppletivismus, může dojít k průniku tvaru do jiného paradigmatu (*incursion*): vysoce autonomní slova/tvary paradigmatu zcela ztratí svůj původní význam, oddělí se od původního paradigmatu a stanou se součástí paradigmatu nového. Např. dnešní komparativ *lepší* byl původně (v praslovanštině) komparativem v paradigmatu s pozitivem *lepý* (Rejzek, 2001, s. 338). Podobně lze uvažovat o dalších paradigmatech (viz § 4.6.3.).⁷

(iv) *Stupeň produktivity konstrukce* je veličina, která je do jisté míry determinována typovou frekvencí. Tato determinace je postavena na předpokladu, že je-li nějaká konstrukce přítomna ve velkém počtu různých jednotek, má tendenci být použita i při tvoření nových jednotek, tedy být produktivní. (Produktivita je vedle toho určována i dalšími faktory, např. sémantickou a morfonologickou složitostí.)⁸ Vysvětlení je následující: vyšší typová frekvence dává konstrukci kognitivně silnější reprezentaci, a ta je tak lépe přístupná při tvoření nových jednotek.

„Čím více existuje tvarů, které nesou daný afix, tím je silnější reprezentace tohoto afixu. Čím silnější je reprezentace afixu, tím snadnější k němu bude [kognitivní/mentální – pozn. JK] přístup, když bude nové slovo vyžadovat flexi, a tím větší bude pravděpodobnost, že tento afix bude produktivní.“ (Bybee, 1995, cit. podle Bybee, 2007a, s. 174, vlastní překlad)

(Bybee, 2007b, s. 11, vlastní překlad). Tedy zjednodušeně: redukce je důsledkem technické automatizace při produkci. Takové vysvětlení je však podle Haspelmatha (2008c) nesprávné. Haspelmath redukční efekt vysvětluje predikovatelností s ohledem na posluchače: „...automatizace se často vyskytuje společně s redukcí formy, protože formy, které jsou automatizované pro mluvčího, jsou často také predikovatelné pro posluchače. Ale v těchto případech příčinou redukce není automatizace, nýbrž tendence mluvčího uspořít energii, pokud je část sdělení predikovatelná. Když není automatizovaná forma predikovatelná (např. když diktuji někomu své telefonní číslo), redukce neproběhne“ (Haspelmath, 2008c, s. 59, vlastní překlad). Haspelmath v tomto odkazuje již na Zipfa (1929), jeho funkční vysvětlení je blízké i Pražské škole. – Navíc dále dovozují, že Haspelmathovo vysvětlení má širší záběr – predikovatelnost a tendence uspořít energii totiž vysvětluje také obecnou diachronní tendenci k preferenci kratších frekventovaných tvarů (při konkurenci s funkčně stejnými, ale delšími tvary) i v případech, kdy se ke starším redukovaným tvarům podle analogie s pravidelnými tvary vytvořily dubletní delší tvary (z jiné perspektivy tak toto vysvětlení doplňuje konzervační efekt). Viz komparativ *snazší* v § 4.6.1., odst. (k).

⁷ Nejedná se o jediný zdroj vzniku suppletivismu. Suppletivismus vzniká i dalšími procesy, např. jako důsledek jazykového kontaktu.

⁸ Záleží např. na tom, do jaké míry je daná třída morfonologicky otevřená a jakým způsobem je třída vymezena (viz úvod § 4.6.).

Naopak tvary o vysoké tokenové frekvenci je možné naučit se mechanicky, bez vztahu k dalším tvarům (viz autonomie), a ty proto nepřispívají k produktivitě dané konstrukce (ibid.).

Tokenová i typová frekvence tak mají podle Bybeeové podíl na vytváření kategorií, tedy formování gramatiky. Tokenová frekvence má vliv na percepci centra kategorie i jejích hranic, zatímco vlastnosti typů vytvářejí hranice kategorie a frekvence typů se vztahuje ke stupni produktivity konstrukce odkazující k této kategorii.

Bybeeová se vyjadřuje i k otázce, zda se věnovat příčinám frekvence, nebo jejím důsledkům. Z textu vyplývá, že samotný původ frekvence není v jejím pojetí jazyka (gramatiky) tak důležitý (vnějazykové příčiny mohou být různorodé) jako její důsledky, protože frekvence (opakování) má vliv na kognitivní reprezentace a právě „kognitivní efekty utvářejí gramatiku“ (Bybee, 2007b, s. 18).

1.2. Frekvence jako explanace jazykových asymetrií

Všechny výše uvedené frekvenční efekty byly Bybeeovou vysvětleny na základě kognitivních principů, které zprostředkovávají vztahy mezi zjištěnou frekvencí a gramatikou. Tato teoretická východiska však není nezbytné přijímat, frekvenci tvarů lze zkoumat nezávisle a považovat ji jednoduše za jeden z faktorů, který indikuje, proč určité struktury mají univerzální tendenci vypadat právě tak, jak vypadají. Takovým způsobem přistupuje k frekvenci Haspelmath (2008a, 2008b). Frekvence je u něj měřitelnou funkční explanací podoby jazykových struktur (jejich vysvětlující proměnnou). Haspelmath především ilustruje, jak frekvence dokáže lépe než některé typy jiného explanačního principu, ikonicity, vysvětlovat univerzální (mezijazykové) asymetrie v jazykovém systému.⁹

Pro přesné uchopení tohoto postupu je opět třeba rozlišit dva pojmy: frekvenci absolutní a frekvenci relativní. *A b s o l u t n í f r e k v e n c e* udává absolutní počet výskytů (všechny výše uvedené případy frekvence se týkaly právě tohoto pojmu). Oproti tomu *r e l a t i v n í f r e k v e n c e* vyjadřuje poměr absolutní frekvence jedné kategorie (tvaru) vůči absolutní frekvenci jiné kategorie (tvaru). Např. lze zkoumat relativní frekvenci superlativu vůči komparativu, vztáhnout přitom mohu absolutní počty komparativů a superlativů, ale zároveň mohu změřit zvlášť relativní frekvenci pro každou dvojici tvarů. Lze tak např. porovnávat poměry superlativu vůči komparativu u dvou různých adjektivních lemmat (např. *nejrůznější* : *různější*; *největší* : *větší*), na absolutní tokenové frekvenci jednotlivých tvarů v tomto případě nezáleží.

Haspelmath (2008a) ukazuje, že relativní frekvencí lze vysvětlit a predikovat univerzální asymetrie v tvarových strukturách, a to i ty, které bývají jinak popisovány a vysvětlovány pomocí tří druhů ikonicity. V jiné studii (Haspelmath, 2008b) předkládá silnou hypotézu:

⁹ Tedy nikoliv přímo jazykově specifické asymetrie. Příklad univerzální morfosyntaktické asymetrie: plurálové tvary u substantiv jsou napříč jazyky delší než odpovídající tvary singulárové.

„Všechny univerzální morfosyntaktické asymetrie mohou být vysvětleny na základě frekvenčních asymetrií, tj. všechny vykazují ekonomickou motivaci: frekventovanější vzorce jsou vyjádřeny menším množstvím materiálu.“ (Haspelmath, 2008b, s. 185, vlastní překlad)

K tomu doplňuje, že ekonomická motivace se do jazyků zapracovává v průběhu historického vývoje (tedy historickými změnami). Vztah mezi frekvencí a jazykovými strukturami tak není přímý. Zjednodušeně řečeno: existují univerzální frekvenční asymetrie (např. méně plurálových jmen než singulárových), které mají tendenci se v jazycích proslazovat jako morfosyntaktické asymetrie (plurálový tvar delší než singulárový).

Tři druhy ikonicity (tedy vztahu podobnosti mezi formou a významem), které se při použití frekvenční explanace stávají nadbytečnými, jsou *ikonická kvantita* (větší množství významu je vyjádřeno větším množstvím formy, např. pozitiv – komparativ – superlativ v latině, ale i češtině), *ikonická komplexnost* (složitější významy jsou vyjádřeny složitějšími formami, pro stupňování platí obdobně) a *ikonická koheze*¹⁰ (významy, které jsou si bližší sémanticky, jsou vyjádřeny koheznějšími formami).¹¹

Ve svém výzkumu na českém materiálu se zaměřím na ikonicitu kvantity a na to, jak koreluje s frekvencí. Ikonická kvantita sice již v současném výzkumu není tak široce přijímána, má ovšem dlouhou tradici a lze ji dobře ilustrovat právě na stupňování adjektiv. Tento koncept lze nalézt v textech Romana Jakobsona, který přímo uvádí:

„Morfologie je bohatá na substitutivní znaky, u nichž je ekvivalentní relace mezi označujícím a označovaným. Tak například v různých indoevropských jazycích stupně srovnání u adjektiv – pozitiv, komparativ a superlativ představují postupné narůstání počtu fonémů; např. angl. *high – higher – highest*, lat. *altus – altior – altissimus*. A tak signifikanty (označující) obřezají gradaci signifikátů (označovaných).“ (Jakobson, 1970, s. 37)

Právě na adresu tohoto příkladu Haspelmath (2008a) uvádí, že je vhodné ho nahradit následující frekvenční explanací: komparativ a superlativ jsou méně frekventované, a proto jsou delší. Podobně přistupuje i ke zbylým dvěma typům: podle něj pro ikonicitu kvantity, komplexnosti a koheze po řadě platí, že kratší, jednodušší a koheznější formy se vyskytují signifikantně častěji než odpovídající delší, složitější a méně kohezní formy. Navíc dále ukazuje, že na rozdíl od ikonicity, frekvenční explanace nevytvářejí nesprávné predikce.

Ve vztahu Haspelmathova přístupu k vlastnímu průzkumu podotýkám následující: asymetrie v systému českého stupňování jsou predikovatelné na základě frekvence, české jazykově specifické asymetrie mohou univerzální asymetrie podpořit, ovšem specifické asymetrie nemohou být frekvencí beze zbytku vysvětleny. Síla Haspelmathových explanací tkví v jejich univerzálním a diachronním charakteru (tzn. prosazují se

¹⁰ *Ikonická koheze* je Haspelmathův (2008a) termín, obvykle bývá označována jako ikonická vzdálenost (*distance*, viz Haiman, 1983). Jde o to, že výrazy, které jsou si bližší konceptuálně, jsou si blíže i na škále koheznosti (vyjádření funkčním slovem – juxtaopozice – vázané formy – portmanteau; viz Haiman, 1983). Haspelmath termín koheze zavádí jako přesnější termín, aby jej odlišil od soumeznosti (*contiguity*).

¹¹ Vedle toho však Haspelmath upozorňuje i na čtyři další typy ikonicity (syntagmatického izomorfismu, soumeznosti, pořadí a opakování), jejichž použitelnost pro lingvistickou teorii nenapadá.

v průběhu jazykového vývoje), nikoliv v jejich okamžité aplikaci na každý specifický jazykový jev. O nastíněných otázkách a hypotézách kolem frekvence a ikonicity pojednám v § 4.2.

2. Metodika práce s korpusem

Vzhledem k tomu, že analýza prostřednictvím frekvence je do velké míry závislá na dobrém zpracování dat z korpusu a hrozí riziko, že bych mohl data omylem nesprávně interpretovat, představím nejprve celý svůj metodický postup. Základním předpokladem pro budoucí rozbor bylo nalézt, pokud možno automaticky, absolutní frekvence výskytu pozitivu, komparativu a superlativu pro všechna adjektiva, která se vyskytují v českých psaných korpusech, a výsledky převést do podoby jednotabulkové databáze. Vzhledem ke snaze získat co nejvíce dokladů rozhodl jsem se prohledávat komplexní korpus SYN, jehož data v době vyhledávání a zpracovávání (říjen 2009 – červen 2010) sestávala z reprezentativních korpusů SYN2000 (100 mil. slov) a SYN2005 (100 mil. slov) a korpusu psané publicistiky z let 1989–2004 SYN2006PUB (300 mil. slov). Učinil jsem tak i s vědomím, že korpus SYN2006PUB „si v žádném ohledu nečiní nárok na reprezentativnost“ (ÚČNK, 2010). Porovnáváním pořadí frekvence komparativů a superlativů v korpusech SYN2000 a SYN2005 oproti korpusu SYN2006PUB jsem totiž zjistil, že dochází jen k takovým dílčím odchylkám, které nemají na celkovou zamýšlenou analýzu vliv. Tyto rozdíly jsou naopak vyváženy výrazně větším množstvím doložených lemmat. Vážnější odchylku jsem zjistil pouze v poměru celkové absolutní frekvence komparativu vůči superlativu (tento rozdíl bude objasněn v § 4.1.).

Nejdříve jsem zaměřil pozornost na vyhledání všech syntetických tvarů stupňování adjektiv.¹² V programu Bonito jsem za pomoci řádkového příkazu [tag="A.....IA.*"] nejprve našel všechny výskyty adjektivních tvarů v pozitivu. Z tohoto seznamu jsem následně nechal prostřednictvím programové funkce „frekvenční distribuce“ vypočítat a zobrazit frekvenci pozitivu u každého adjektivního lemmatu (při nastavení, že velikost písmen bude ignorována). Výsledkem byla tabulka o dvou sloupcích, v níž první sloupec obsahoval označení lemmatu (základní tvar lemmatu je vždy adjektivum v pozitivu) a druhý danou frekvenci. Tento výpis jsem exportoval a následně převedl do tabulkového editoru. Totožný postup jsem opakoval s komparativem a superlativem (s náležitou záměnou číslovky 1 za 2, resp. 3 ve formulí).

¹² Přesněji řečeno: (i) Zaměřil jsem se na všechny jednotky, které vykazují formální morfologické vlastnosti adjektiv se složenou deklinací. Do zkoumání tak není zahrnuto jmenné adjektivum/adverbium *rád*. Ačkoliv *rád* může stát ve frekventované syntaktické pozici adjektiva v predikativu (což platí i pro stupňované nesklonné tvary (*nej-raději, raděj, radši, radčí, rači, ...*), jak upozornil jeden z recenzentů tohoto textu), jeho tvary jsou morfologicky relevantní pro výzkum adverbialního stupňování. (ii) Nezahrnul jsem do vyhledávání automaticky označené záporné tvary adjektiv (k tomu viz § 2.1.). (iii) Neprohledával jsem tvary stupňování analytického (zejm. typu (*nej-více* + adjektivum)). Na základě orientačního průzkumu se jejich celková tokenová frekvence ukázala jako velmi nízká (ve srovnání se syntetickým stupňováním méně než setinová). Vzhledem k definovaným cílům jsem roli analytického stupňování vyhodnotil jako zanedbatelnou. Detailní výzkum těchto konstrukcí by však mohl být užitečný z hlediska jejich rozložení mezi lemmata i z hlediska analýzy typů syntaktických vztahů, v nichž se prosazují.

Takto vytvořené tři tabulky jsem v několika krocích sloučil v jedinou databázi, která měla pro každé lemma na řádku ve třech sloupcích frekvenci pozitivu f (POZ), komparativu f (KOMP) i superlativu f (SUP) (včetně nulových hodnot). Získal jsem tak z korpusu SYN vyčerpávající soubor všech automaticky označených adjektiv s frekvencemi jejich stupňovaných tvarů.

Následně jsem provedl hrubou analýzou dat a jejich třídění. Nejprve jsem se rozhodl vyřadit z dalšího výzkumu ta adjektiva, která neměla jediný výskyt komparativu ani superlativu. Z původního seznamu 65 400 lexémů tak zůstalo přibližně 5 400 stupňovaných adjektiv.¹³

Tabulku jsem dále doplnil o sloupce, které zachycují pro daná lemmata další zajímavé hodnoty: (a) součet frekvencí všech tří stupňů; (b) součet frekvencí komparativu a superlativu – dále $f(\kappa+s)$; (c) relativní frekvence komparativu vůči součtu frekvencí komparativu a superlativu; relativní frekvence (d) pozitivu, resp. (e) komparativu, resp. (f) superlativu vůči součtu frekvencí všech tří stupňů; (g) relativní frekvence součtu komparativu a superlativu vůči součtu frekvencí všech tří stupňů. Tabulku jsem se poté rozhodl seřadit sestupně podle hodnot $f(\kappa+s)$ ¹⁴ a data jsem začal detailněji kontrolovat, popisovat a následně analyzovat.

2.1. Zpracování agregovaných dat: kvalitativní kontrola a ruční značkování lexémů

Podrobně jsem zkontroloval a označoval prvních 500 adjektiv s nejvyšší $f(\kappa+s)$. Pro jednotlivá „podezřelá“ lemmata jsem ověřil, jaké tvary jejich KOMP a SUP byly ve skutečnosti označovány a jaké naopak chybějí. Dále jsem ručně vypsal u všech položek tvar komparativu, a pokud jsem předpokládal, že by u daného lemmatu mohla existovat více než jedna řada tvarů KOMP a SUP, ověřoval jsem v korpusu zastoupení všech teoreticky možných tvarů (včetně neobvyklých tvarů redukováných i neredukovaných: *tvrdější, krasší, sladčí* apod.). V některých případech byly již frekvence započítány (např. všechny stupňované tvary *snadnější, snazší* i *snazší*), v jiných nikoliv (*zazší* a *nejzazší* nebylo přiřazeno k *zadnější* a *nejzadnější*).¹⁵

V tabulce jsem pak u lemmat, která vykazovala dubletní stupňování, rozlišil frekvenci hlavního tvaru KOMP a SUP (= frekventovanějšího) a vedlejšího tvaru (= méně frekventovaného).¹⁶ Původní základ tabulky lemma, f (POZ), f (KOMP), f (SUP) jsem rozšířil o f (hlavní KOMP), f (hl. SUP), f (vedlejší KOMP), f (vedl. SUP). Kromě toho jsem pro snazší orientaci označil ve zvláštním sloupci křížkem lemmata s dvojími stupňovanými tvary. U jednotlivých lemmat jsem dále do zvláštního sloupce binárně rozlišil, zda je

¹³ Po dalších úpravách a dohledáváních se počet nakonec ustálil na 5 440 lemmatech.

¹⁴ Za tímto rozhodnutím stála jednak snaha neopomenout při další práci lemmata, jejichž jeden ze stupňovaných tvarů je málo frekventovaný, jednak představa, že frekvence komparativů i superlativů mohou vysvětlovat morfemickou strukturu stupňovaných tvarů stejnou silou.

¹⁵ Je ovšem třeba podotknout, že v rámci korpusu SYN je značkování na rozdíl od staršího korpusu SYN2000 (v němž jsem původně začal vyhledávat) ve velké míře spolehlivé; korekce u frekventovanějších adjektiv byly minimální.

¹⁶ V několika izolovaných případech bylo více různých vedlejších tvarů, viz celková analýza v § 4.6.1.

stupňování neredukované (komparativ/superlativ v korpusu doložený pouze s hláskovou změnou na konci kmene a příponou *-ejší*), nebo redukováné (alespoň jeden doložený stupňovaný tvar kratší než při pravidelném tvoření).¹⁷ Účelnost tohoto rozlišování se ukáže později.

Následně jsem zjednodušeným způsobem prošel i dalších 1 000 položek (podle stejných pravidel, ovšem bez ručního vypisování komparativů) a nakonec v celém souboru zbylých 3 900 slov dohledal podle všech „podezřelých“ zakončení pozitivu (*-bý*, *-dký*, *-dý* apod.) slova, u nichž jsem předpokládal, že by mohla mít nepravidelné nebo dubletní stupňování. Tato zakončení jsem dále dohledával i v korpusu SYN, abych doplnil případné neoznačkové redukováné komparativy. Odtud byl už jen krok k tomu, abych mohl sestavit inventář všech zakončení s frekvencemi (v § 4.6.1.).

Z pořízených dat jsem dále vyloučil zvláštní superlativní kategorii slov s morfematickou strukturou *sebe* + KOMP, tedy tvary jako *sebemenší*, *sebelepší*. Tyto tvary byly v automaticky pořízených datech zobrazeny nesystematicky. V korpusu SYN jsem proto tyto tvary s frekvencemi vyhledal zvlášť pomocí tagu [lc="sebe.*ší.*"], následně jsem jejich frekvence exportoval, prověřil a uložil do zvláštní tabulky. K předběžné analýze těchto tvarů viz § 3.1. V databázi jsem naopak ponechal data, v nichž má morfém *sebe* odlišnou derivační historii, tedy např. *sebevědomý*.

Jak jsem uvedl v pozn. 12, součástí celkové analýzy nejsou automaticky označené záporné stupňované tvary adjektiv, a to ze tří důvodů. (i) Výskyt těchto tvarů je až na několik výjimek (*nebezpečnější*, *nepříjemnější*, *nemenší*) frekvenčně silně omezený. (ii) Průzkum jednotlivých kontextů záporných tvarů ukázal, že u značného počtu lemmat jsou veškeré domnělé doklady pouze chybně napsanými tvary superlativu (s elizí *j* v předponě *nej-*; týká se to mj. řetězců *nevyšší*, *nevětší*, *nebližší*). Zahnutí těchto tvarů do analýzy by proto bez důkladné kontroly kontextů všech výskytů u 634 lemmat nebylo užitečné (skutečný počet lemmat se bude pravděpodobně blížit 276, což je počet záporných lemmat v superlativu *nejne-*). (iii) Záporné tvary jsou dvojího typu: jednak vlastní stupňované tvary záporných adjektiv (*nebezpečný* > *nebezpečnější*), jednak negované tvary komparativu (*menší* > *nemenší*, nikoliv tedy *nemalý* > *nemenší*). Kvůli všem těmto omezením jsem se nakonec rozhodl ručně dohledat výskyty těchto tvarů jen u lemmat s nepravidelným stupňováním (viz § 4.6.1.).

Na základě průzkumu, který jsem činil průběžně při tomto zpracování dat, volím pro vlastní analýzu následující postup: nejprve zformuluji hypotézy týkající se morfologického vyjádření (morfologických markerů) stupňování a frekvence, dále představím základní strukturní a frekvenční charakteristiky celého souboru a následně budu dané hypotézy testovat na sebraném materiálu a souběžně s tím s jejich pomocí vybuduji vyčerpávající klasifikaci syntetických tvarů českého stupňování adjektiv.

¹⁷ Viz výše pozn. 6 o platnosti redukčního efektu. – Za redukováné považuji pro účely této práce všechny případy, kdy je počet hlásek v komparativu nižší než při pravidelném tvoření příponou *-ejší*. Nepovažuji tak za redukci hláskovou alternaci *r – ř*: *přikřejší – přikrejší* apod.).

3. Hypotézy o vztazích frekvence a morfologického značení; základní charakteristiky souboru

Na základě výše uvedených teoretických poznatků o vztahu frekvence a morfologického vyjádření předpokládám, že získaná korpusová data potvrdí, či nepotvrdí následující hypotézy:

- (1) **Redukční efekt celoskupinový:** Redukované tvary¹⁸ komparativu a superlativu budou soustředěny u lemmat s vysokou frekvencí tvarů komparativu a/nebo superlativu. Jinými slovy: Čím více budou komparativní a/nebo superlativní tvary frekventované, tím vyšší je pravděpodobnost, že budou kratší (oproti alternativním tvarům pravidelným).
- (2) **Efekt autonomie celoskupinový:** K supletivismu bude docházet u lemmat s nejvyšší frekvencí tvarů komparativu a/nebo superlativu.
- (3) **Redukční efekt vnitroskupinový:** U některých lemmat existují vedle redukovaných stupňovaných tvarů i neredukované tvary. Porovnáme-li dvě taková lemmata v rámci jednoho typu z hlediska relativní frekvence jejich redukovaných a neredukovaných tvarů, předpokládám, že u lemmat s vyšší absolutní frekvencí tvarů komparativu a/nebo superlativu bude poměr redukovaných tvarů vůči neredukovaným vyšší než u lemmat s nižší absolutní frekvencí.
- (4) **Konzervační efekt:** Pokud dochází v rámci skupiny lemmat se stejným kmenovým zakončením ke zpravidelňování stupňování, lemmata s vysokou frekvencí tvarů komparativu a/nebo superlativu mu nepodléhají nebo mu podléhají méně než lemmata s nízkou frekvencí.
- (5) **Efekt produktivity:** Lemmata stupňovatelná příponou *-ejší*, která se zdá splňovat ostatní znaky produktivity, budou nejčetnější z hlediska typové frekvence.
- (6) **Kvantita a morfosyntaktická asymetrie:** Superlativní tvary, které jsou delší a složitější než komparativní tvary, budou celkově méně frekventované než komparativní tvary. Data v češtině tak mohou podpořit univerzální morfosyntaktickou asymetrii.

Poslední dva případy nejsou hypotézami se stejnou platností jako šest výše uvedených. Hypotéza (7) může vysvětlit, proč se některá lemmata vymykají obecně platnému schématu. Efekt autonomie (8) nelze jako hypotézu zvoleným způsobem potvrdit, přijmu-li však dané koreláty, mohu s jeho pomocí na základě frekvence načrtnout těsnost vztahů v rámci českého stupňování.

(7) **Kvantita a relativní frekvence v rámci lemmatu:** U některých lemmat může nastat ojedinělá asymetrie, že výrazně frekventovanější superlativní tvar bude kratší než méně frekventovaný komparativ.

¹⁸ Všechny obecně krátké tvary, ne nutně vzniklé redukčním procesem, viz výše pozn. 6 a 17.

(8) Efekt autonomie vnitrolemmatový: Frekventovanější tvary budou autonomnější. V rámci trojice pozitiv – komparativ – superlativ nejfrekventovanější z tvarů bude nejautonomnější a v rámci gramatiky/slovníku bude mít slabší vztahy ke dvěma zbylým stupňům v lemmatu. Autonomnější budou ty nejfrekventovanější tvary.

3.1. Základní strukturní a frekvenční charakteristiky

Při ověřování hypotéz budu pracovat s různými frekvenčními charakteristikami. Základní data souboru představím v tomto oddíle.

(i) Údajem relevantním zejména pro redukční efekt je celkový počet lemmat s redukovaným stupňováním:¹⁹ redukovaný tvar komparativu/superlativu je doložen celkem u 72 lemmat z celkového počtu 5 440 adjektivních lemmat (jde tedy o celkový podíl 1,3 %).

(ii) Dále musím rozhodnout, která frekvenční charakteristika (zda frekvence komparativu, superlativu, nebo všech stupňovaných tvarů) má zásadní význam pro morfosyntaktickou strukturu. Dosud jsem předpokládal, že frekvenční efekty může vysvětlovat frekvence komparativních tvarů stejnou silou jako frekvence superlativních tvarů. Je však na místě zvážit a porovnat, zda samotná frekvence komparativu, nebo superlativu nemůže mít vyšší explanační sílu. Za tímto účelem jsem provedl porovnání frekvence tvarů u redukovaných lemmat:

Obecně platí, že podíl lemmat s redukovaným stupňováním (na všech lemmatech) klesá s klesající tokenovou frekvencí. Databáze vykazuje při seřazení podle $f(\kappa+s)$ rozložení takové, že v prvních 51 lemmatech je soustředěno celkem 26 redukovaných lemmat. To znamená, že zatímco celkový podíl redukovaných lemmat je 1,3 % (viz výše), omezím-li výběr na skupinu 51 lemmat s nejvyšším součtem tvarů komparativu a superlativu, poměr odpovídá 51 %. U dalších 5 389 lemmat je redukce doložena jen v 46 případech. (I těchto zbylých 46 lemmat je rozloženo asymetricky, jejich výskyt významně ubývá s klesající $f(\kappa+s)$.)

Při srovnávání frekvenčních charakteristik se dále ukázalo, že frekvence samotného komparativu indikuje redukci lépe než $f(\kappa+s)$: výsledky podle $f(\text{KOMP})$ vykazují rozložení 23 redukovaných lemmat mezi 34 nejfrekventovanějšími komparativy. Naopak frekvence samotného superlativu $f(\text{SUP})$ je výrazně horší indikátor. Rozdíly ilustruji v § 4.1.; indikují je i seznamy 25 nejfrekventovanějších lemmat podle každé z charakteristik (nejvíce redukovaných lemmat je koncentrováno na prvních místech podle $f(\text{KOMP})$):

- (a) 25 nejčastějších lemmat, seřazeno podle $f(\kappa+s)$: *velký, dobrý, vysoký, malý, starý, nízký, špatný, blízký, mladý, dlouhý, důležitý, silný, různý, významný, široký, levný, drahý, nový, rychlý, těžký, pozdní, slabý, brzký, úspěšný, známý.*

¹⁹ Lemma s redukovaným stupňováním = lemma, jehož alespoň jeden doložený stupňovaný tvar je redukovaný.

- (b) 25 nejčastějších lemmat, seřazeno podle f (KOMP): *velký, vysoký, dobrý, malý, starý, nízký, mladý, dlouhý, špatný, široký, pozdní, brzký, silný, levný, drahý, slabý, blízký, rychlý, důležitý, těžký, krátký, složitý, výrazný, hluboký, výhodný*.
- (c) 25 nejčastějších lemmat, seřazeno podle f (SUP): *velký, dobrý, vysoký, blízký, důležitý, různý, špatný, starý, malý, významný, nový, silný, nízký, úspěšný, známý, mladý, krásný, těžký, slavný, rychlý, častý, drahý, oblíbený, levný, zajímavý*.

Rovněž při porovnání větších částí souboru se ukázalo, že většina lemmat s redukovánými tvary zaznamenala při seřazení podle f (KOMP) oproti seřazení podle f ($\kappa+s$) posun v pořadí výše (opačně se posunulo pouze několik lemmat, nejvýrazněji lemma *zadní* se superlativem *nejzazší*), zatímco lemmata s neredukovanými vysoce frekventovanými tvary superlativu a relativně nižší frekvencí tvarů komparativu (*nejdůležitější, nejruznější, nejvýznamnější, nejúspěšnější, nejznámější, nejslavnější, nejčastější, nejoblíbenější* atd.) se posunula níže. Hodnoty absolutní frekvence u redukováných tvarů tedy naznačují, že f ($\kappa+s$) – vyjma několika izolovaných případů – pouze snižuje rozdíly mezi indikátorem lepším f (KOMP) a horším f (SUP).

Zřetelně se ukázalo, že frekvence samotného komparativu má ze tří sledovaných charakteristik nejvyšší explanační sílu. Klesající exponenciální průběh podílu redukováných lemmat v závislosti na frekvenci (včetně srovnání tří charakteristik) je zaznamenán v grafu na obr. 1 v § 4.1.

V dalším textu již budu primárně uvažovat seřazení podle frekvence komparativních tvarů f (KOMP). V tomto smyslu přeformulovávám i jednotlivé hypotézy a efekty: místo „frekvence tvarů komparativu a/nebo superlativu“ měním v hypotézách (1), (2), (3), (4) všechny formulace na „frekvence tvarů komparativu“.

(iii) Tvoření superlativu, struktura komparativu a superlativu a celkový vztah frekvence mezi komparativem a superlativem. Strukturně je zřejmé, že superlativ se tvoří na rozdíl od komparativu zcela pravidelně. O formální struktuře českého stupňování tak lze říci následující: Zatímco český komparativ se tvoří, za obvyklé účasti hláskových změn v kmeni, jak pravidelným sufixem, tak dalšími redukovánými způsoby na základě tvarů pozitivu i supletivně, české superlativní tvary jsou složeny vždy z neredukovatelného prefixu *nej-*, který se zcela pravidelně pojí s tvarem komparativu.

Vedle klasického superlativu se na základě tvaru komparativu tvoří i zvláštní superlativní tvary s prefixem *sebe-*. Jejich distribuce je v korpusu velmi omezená; v korpusu SYN se jedná o celkem 11 217 dokladů u 816 lemmat, z toho celých 53,7 % (!) dokladů patří k tvaru *sebemenší*. Frekvence deseti nejčtenějších tvarů ukazuje opět asymetrické rozložení: *sebemenší* (6 029), *sebelepší* (1 186), *sebevětší* (634), *sebedokonalejší* (178), *sebenepatrnější* (159), *sebekrásnější* (107), *sebehorší* (52), *sebekratší* (46), *sebedelší* (44). Jak indikuje i pátý nejčtenější tvar *sebenepatrnější*, *sebe-*superlativy se užívají ve specifických, obvykle sémanticky negativně zabarvených kontextech. Jejich význam se popisuje např. jako „připuštění nejvyšší míry vlastnosti“ (Karlík – Nekula – Rusínová, 2003, s. 178) nebo „marnost“ (Marvan, 2008, s. 69). Tvary této třídy se stejně jako tvary *nej-*superlativu tvoří zcela pravidelně.

Pro ověření hypotézy kvantity je základním ukazatelem celková absolutní frekvence komparativních a superlativních tvarů. Z celkového počtu 2,6 mil. stupňovaných

tvarů v korpusu SYN je 1,38 mil. komparativních (53,17 %) a 1,22 mil. superlativních tvarů (46,83 %).

(iv) Frekvenční vztahy mezi pozitivem, komparativem a superlativem. Celková frekvence tvarů pozitivu podle očekávání výrazně překračuje frekvence stupňovaných tvarů. V rámci lemmat s doloženým stupňováním odpovídá 2,6 mil. stupňovaných tvarů 32,8 mil. tvarů pozitivu. Podíl stupňovaných tvarů je obecně vyšší u lemmat s vysokou frekvencí komparativu a superlativu a dále také u několika výjimečných případů, viz pozn. 22.

(v) Vztah frekvence přípony *-ejší* a dalších způsobů tvoření komparativu. Přípona *-ejší* se vyskytuje jako jediný způsob tvoření v 5 369 lemmatech a dále dubletně v dalších 24 lemmatech; její celková tokenová frekvence činí 475 182 komparativů a 475 824 superlativů. Ostatní způsoby tvoření se oproti tomu vyskytují celkem pouze v 71 lemmatech; celková tokenová frekvence ostatních sufixů činí 879 262 komparativů a 728 172 superlativů.

4. Ověření hypotéz

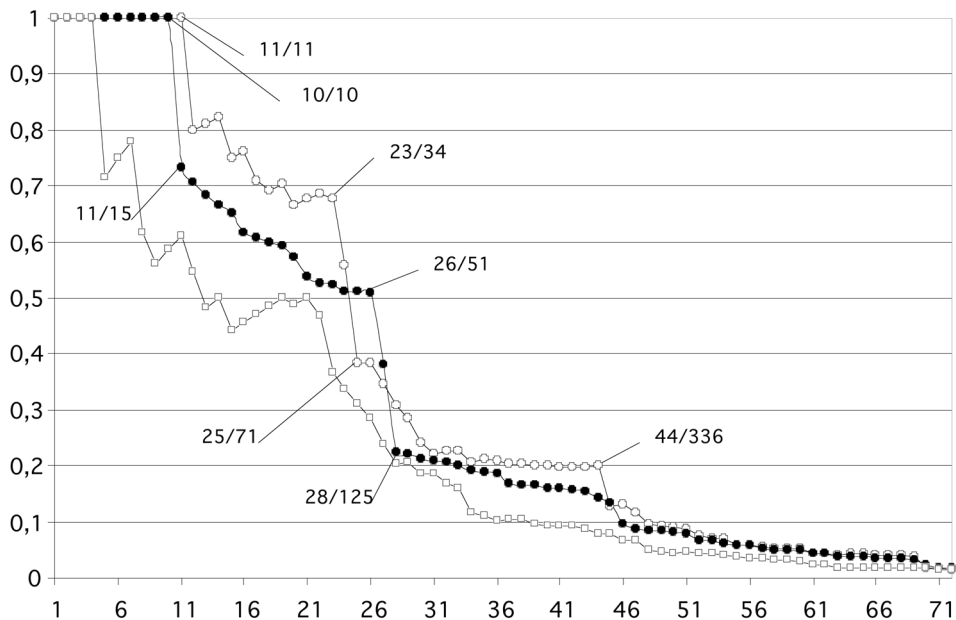
4.1. Rozložení redukovaných a supletivních tvarů: hypotézy (1) a (2)

Celokupinový redukční efekt (1) lze nejlépe odhalit podrobným prozkoumáním frekvence všech redukovaných tvarů v databázi. Učinil jsem to již v § 3.1., kde jsem ukázal, že klesající tendenci podílu redukovaných lemmat nejlépe vysvětluje frekvence samotného komparativu. Z hlediska efektu autonomie (2) pak pozoruji, že redukované nebo supletivní stupňování má celkem jedenáct adjektiv s nejvyšší frekvencí tvarů komparativu (*velký, vysoký, dobrý, malý, starý, nízký, mladý, dlouhý, špatný, široký a pozdní*). Obě pozorování jsou tedy v souladu s očekáváním. Přistoupím nyní k podrobnějšímu prověření obou hypotéz.

Ověření celokupinového redukčního efektu (1). Rozložení redukovaných tvarů je koncentrováno mezi lemmaty s nejméně frekventovanějšími komparativy; výskyt redukovaných lemmat se signifikantně snižuje s klesající f (KOMP): seřadím-li lemmata podle klesající f (KOMP) a změřím-li pro každé redukované lemma podíl jeho pořadí v souboru redukovaných vůči jeho pořadí v souboru všech lemmat, nalézám v souboru redukovaných jednoznačnou korelaci: s každým dalším redukovaným tvarem tento podíl klesá. Exponenciální regresní model vystihuje klesání s koeficientem determinace 97,2 %. Tato statistika kvantifikuje zřejmý průzkum, který jsem provedl v § 3.1. Stav zachycuje graf na obr. 1. Graf také znázorňuje, jak f (KOMP) vysvětluje redukci stupňovaných tvarů v porovnání s f (K+S) a f (SUP).

Ověření celokupinového efektu autonomie (2). Všechny komparativy supletivních lemmat (především *lepší, horší, dřívější*; a také *větší, menší, delší*, příp. *pozdější*) mají velmi vysokou frekvenci (patří mezi 12 nejčastějších komparativů), což potvrzuje celokupinový efekt autonomie.

Uvedené poznatky o superlativu a komparativu pak vyzývají k odpovědi na další otázky: jaké jsou obecné vztahy mezi frekvencí (strukturou) komparativu na jedné



Obr. 1: Podíl redukovaných lemmat na celkovém počtu lemmat klesá s $f(KOMP)$, $f(K+S)$ i $f(SUP)$. Na ose x je značeno pořadí redukovaného lemmatu mezi redukovanými; osa y značí podíl pořadí redukovaného lemmatu mezi redukovanými k jeho pořadí mezi všemi lemmaty. \circ = pořadí podle klesající $f(KOMP)$, \bullet = pořadí podle klesající $f(K+S)$, \square = pořadí podle klesající $f(SUP)$.

straně a superlativu na straně druhé? V těchto souvislostech prozkoumám hypotézu kvantity (6), otázku specifické asymetrie (7) a efekt autonomie (8).

4.2. Délka stupňovaných tvarů: hypotéza (6)

Na základě univerzální morfosyntaktické asymetrie (6) jsem predikoval, že superlativní tvary v češtině budou celkově méně frekventované než tvary komparativní. V § 3.1. jsem uvedl, že superlativ se (na rozdíl od komparativu) tvoří zcela pravidelně na základě tvaru komparativu. Přistoupím teď k podrobnému prověření frekvenčních charakteristik.

Ověření kvantity a morfosyntaktické asymetrie $KOMP$ a SUP (6). Je možné obecně delší tvar superlativu vysvětlit nižší frekvencí tvarů superlativu v porovnání s komparativem? Odpověď není na první pohled zcela jednoznačná: (i) Existuje mnoho lemmat, která mají více tvarů superlativu než komparativu. (ii) V § 3.1. jsem uvedl, že z celkového počtu 2,6 mil. stupňovaných tvarů v korpusu SYN je 53,17 % komparativů a 46,83 % superlativů. Takový rozdíl mezi počty $KOMP$ a SUP je sice statisticky signifikantní (alternativní hypotéza, že počet superlativů je stejný nebo vyšší než počet komparativů, je v 99% intervalu spolehlivosti vyloučena), ale v absolutních číslech nepatrný.

V této souvislosti jsem také ověřil, zda nemůže mít na celkovou frekvenci vliv zařazení rozsáhlého publicistického korpusu SYN2006PUB. Ukázalo se, že tento vliv není bezvýznamný: v publicistických textech se užívá celkově více superlativů v porovnání s komparativy než v jiných žánrech. Absolutní rozdíl mezi KOMP a SUP je tak ve vyváženém korpusu větší. Porovnání uvádí tab. 1.

Tab. 1: Celková relativní frekvence KOMP a SUP

	SYN	SYN2000	SYN2005	SYN2006PUB
abs. podíl KOMP	53,17 %	55,0 %	56,5 %	51,6 %
abs. podíl SUP	46,83 %	45,0 %	43,5 %	48,4 %

Navíc (iii) podle kritéria relativní frekvence (porovnání frekvencí jednotlivých lemmat), což je metoda, která je pro kategoriální porovnání KOMP a SUP adekvátnější, je průměrná relativní frekvence komparativu jednoznačně vyšší: v korpusu SYN 64,21 %. Zdůrazňuji, že při tomto způsobu výpočtu jsem vypočítal průměr z poměrů u jednotlivých lemmat, stejnou váhu tak má lemma nejfrekventovanějšího komparativu jako lemma nejméně frekventovaného komparativu.²⁰ Zde se naopak mírně vymyká korpus SYN2005, SYN2006PUB má nejvyšší průměrnou relativní frekvenci komparativu (viz tab. 2).

Tab. 2: Průměrná relativní frekvence KOMP a SUP jednotlivých lemmat

	SYN	SYN2000	SYN2005	SYN2006PUB
rel. podíl KOMP	64,21 %	64,13 %	62,91 %	64,35 %
rel. podíl SUP	35,79 %	35,87 %	37,09 %	35,65 %

Nakonec jsem předběžně²¹ porovnal obě kategorie podle toho, v kolika různých lemmatech převažuje komparativ nad superlativem a naopak. Ukázalo se, že u více než dvou třetin lemmat je relativní frekvence komparativu vyšší než superlativum.

P o d p o r e n í e f e k t u p r o d u k t i v i t y (5). V této souvislosti připomínám pojem autonomie a stanoviska Bybeeové (§ 1.1.). Podle nich lze dovodit, že z hlediska produktivity nejsou důležité vlastnosti několika nejfrekventovanějších lemmat. Největší vliv na kognitivní utváření kategorie a produktivitu mají právě ty stupňované tvary, které mají nižší tokenovou frekvenci a zároveň vysokou typovou frekvenci (tzn. ty se společnou frekventovanou formální charakteristikou), jde tedy o lemmata, u nichž převažuje komparativ nad superlativem.

Shrnutí: Hypotéza kvantity se v datech potvrzuje, a to především podle kritéria relativní typové frekvence. Data ukazují, že v češtině existuje více lemmat, u nichž převa-

²⁰ Rozdíly mezi celkovou a průměrnou relativní frekvencí mj. vypovídají o tom, že vysoký podíl superlativů se vyskytuje u lemmat s vyšší absolutní frekvencí stupňovaných tvarů (viz např. vyšší počet SUP než KOMP u lemmat *velký, dobrý*), zatímco méně frekventované stupňované tvary mají vyšší podíl komparativů (v krajním případě se to projevuje tak, že v SYN je 1 753 lemmat s doloženým KOMP bez SUP, zatímco jen 806 lemmat s opačnou charakteristikou).

²¹ Tyto výsledky je nutné považovat za orientační. Pro správné vztahování bych musel pro každé lemma ověřit, zda rozdíl mezi počty komparativů a superlativů je statisticky signifikantní.

žuje komparativ nad superlativem. Přitom se jedná o lemmata s nižší frekvencí stupňovaných tvarů, a ta na rozdíl od těch nejfrekventovanějších přispívají k produktivitě kategorie.

4.3. Kvantita a relativní frekvence v rámci lemmatu: hypotéza (7)

O v ě ř e n í. Jak jsem již výše uvedl, superlativ se v češtině tvoří zcela pravidelně, v té souvislosti se zdá krajně nepravděpodobné, že v češtině existují tvary superlativu kratší než tvary komparativu. Jediným případem, který se tomu blíží, je již jednou zmíněné lemma *zadní*. Ačkoliv jsou doloženy kratší i delší tvary pro oba stupně, v komparativu se častěji používá delší tvar a v superlativu varianta kratší. Tuto asymetrii (*zadnější : nejzazší*), která historicky vznikla jako spojení dvou různých řad tvarů, je možné vysvětlit právě rozdílem v relativní frekvenci KOMP : SUP oproti jiným lemmatům (frekvence tvarů lemmatu *zadní* v korpusu SYN: *zadní* (28 811), *zadnější* (7), *zazší* (2), *nejzadnější* (26), *nejzazší* (1 482)). Superlativ je mnohem častěji používaný, a proto u něj převažuje kratší tvar.

4.4. Vnitrolemmatový efekt autonomie (8)

Podle efektu autonomie platí, že čím je tvar frekventovanější, tím spíše bude mít vlastní reprezentaci. Pokusím se zjistit, zda a jak se tato tendence projevuje nikoliv u jednotlivých tvarů, ale v rámci tříd českého stupňování, tedy v různých stupních autonomie mezi kategoriemi pozitiv – komparativ – superlativ. Vycházím z předpokladu, že mezi třemi kategoriemi existují různě silné gramaticko-lexikální vztahy.

N a z n a č e n í e f e k t u a u t o n o m i e (8). Frekvenčně se jako nejautonomnější obecně ukazují tvary pozitivu. Jejich četnost výrazně převyšuje stupňované tvary, viz výše v § 3.1.²² Jako nejméně autonomní se naopak jeví tvary superlativu: morfoloogicky a sémanticky ve vztahu k nižším stupňům jsou zcela pravidelné (prefix superlativu není redukovatelný). Zároveň se ukazuje, že absolutní frekvence tvarů superlativu ani nepřispívá k redukci sufixové části; tedy i ty nejautonomnější superlativy jsou pravidelné.

Kategorie komparativu pak vykazuje jako celek obojí znaky: ve většině lemmat se ve vztahu k pozitivu jedná o tvary méně autonomní, pravidelné (sémanticky, formálně, z hlediska nižší frekvence), ovšem existuje skupina nejméně 78 nepravidelných komparativů, které se danému schématu vymykají. O tom podrobněji v § 4.6.

4.5. Efekt produktivity (5)

Bližší pohled na skladbu redukovaných tvarů na jedné straně a produktivních tvarů na straně druhé (viz § 3.1.) umožňuje ověřit, zda lze za nejčastější a produktivní gra-

²² Součet stupňovaných tvarů převyšuje počet tvarů pozitivu pouze u 15 lemmat (např. *lidnatý, pozdní, nízký*). Přibližně 50% podíl pozitivu na všech tvarech vykazují supletivní a redukovaná lemmata s vysokou frekvencí komparativu a superlativu (*blízký, vysoký, dobrý, špatný, velký*).

dační typ považovat strukturu s příponou *-ejší*. Z dat vyplynulo, že tvoření touto příponou vykazuje velmi vysokou typovou frekvenci a současně absolutně nižší frekvenci tokenovou. Jedná se tedy o zřejmou ukázkou produktivního tvoření. Dále lze také vyčíst, že celkový počet komparativů a superlativů je u přípony *-ejší* vyrovnaný. Tento na první pohled překvapivý výsledek lze vysvětlit tím, že redukci nepodléhají právě ty vysoce frekventované superlativy, které mají zároveň nižší frekvenci komparativů. Tím se navyšuje absolutní počet superlativů (relativní frekvence však zůstává ve prospěch komparativů, viz § 4.2.).

4.6. *Redukční vnitroskupinový a konzervační efekt: hypotézy (3) a (4); klasifikace tvarů českého stupňování*

Na závěr se budu věnovat efektům, k jejichž ověření mohu přistoupit až po podrobném popisu všech tříd českého stupňování. Zaměřím se proto nejprve na zjištění všech způsobů tvoření českých stupňovaných tvarů a jejich frekvencí; to mi ve výsledku umožní formulovat stanoviska k hypotézám (3) a (4), tedy k tomu, jak frekvence působí na formální strukturu českého stupňování, a to zejména uvnitř jednotlivých tříd s podobnou stavbou.

Třídy tvarů lze definovat dvěma způsoby (podle Bybee, 1995; 2007a, s. 171): buď z hlediska procesu, jímž je komparativ utvářen (např. CH-í > š-ší),²³ nebo z hlediska výsledného stavu, tvaru komparativu (např. -ší). V jazyce je možné rozpoznávat oba typy tříd. Tvoření českého komparativu se zdá být lépe vymezené procesuálně, tedy rozhodl jsem se komparativy klasifikovat na základě tvaru pozitivu a procesu, který vznik komparativu popisuje. Přesto na několika místech poukážu, že u některých tvarů komparativu (např. u zakončení T-ší, D-ší) jsou náznaky pro kategorii vymezenou výsledným stavem.

4.6.1. *Seznam všech tříd způsobů tvoření komparativu*

Jednotlivé třídy uvádím v následující struktuře: nejprve podávám informaci o počtu adjektivních lemmat s daným zakončením, která se vyskytla v korpusu SYN, následuje popis způsobu tvoření a na závěr uvádím tabulku, v níž je na prvním místě uvedeno pořadí v databázi podle *f* (KOMP) (v případě rovnosti je rozhodující pořadí *f* (SUP) a následně *f* (POZ), pořadí těch tvarů, které nebyly zahrnuty do celkové analýzy, je uvedeno v závorce), následuje tvar pozitivu, komparativu a absolutní počty tvarů pozitivu, komparativu a superlativu v daném lemmatu.

Existuje-li více tvarů komparativu/superlativu v rámci jednoho lemmatu, na prvním řádku uvádím frekvenci hlavního tvaru. V klasifikaci eviduji všechny doložené stup-

²³ Na rozdíl od většiny popisů značím morfonologickou strukturu a její změny namísto grafémů pomocí fonémů (foném *ch* obvykle značený *x* značím pro srozumitelnost CH; za základní považuji vždy morfologicky nejsilnější variantu, tzn. bez ohledu na asimilaci znělosti, např. *z* v *užší*), což graficky odlišuji kapitálkami. Hovořím-li někdy ekvivalentně o hláskách, je to tam, kde rozdíl není důležitý. Dále používám ve schématech tyto zkratky: C pro konsonant, V pro vokál.

ňované tvary adjektiv, neusilují přitom o jejich normativní hodnocení (pokud přímo neodkazují na spisovnou normu; viz § 4.6.2.). Hranici mezi nahodilým použitím a existujícím vývojovým procesem totiž vidím v souladu s nastíněnými teoretickými přístupy jako prostupnou. Výběrová interpretace některých tvarů např. jako pravopisné chyby a jejich nezahrnutí do klasifikace by mohly mít za následek zakrytí možných vývojových tendencí. Stojí však za povšimnutí, že mezi dubletními variantami komparativů a superlativů existují napříč třídami tvarové a frekvenční rozdíly (viz dále zejm. pozn. 26).

(a) Adjektiva zakončená na CH-í

Doloženo celkem 7 lemmat (plus jedno záporné). Komparativ se tvoří zcela pravidelně se změnou poslední kmenové hlásky CH > Š příponou -ší [CH-í > Š-ší].²⁴

# KOMP	tvar POZ	tvar KOMP	f (POZ)	f (KOMP)	f (SUP)
26	jednoduchý	jednodušší	51 843	8 069	4 571
183	tichý	tíšší	25 720	487	212
200	suchý	sušší	25 689	436	152
501	ploché	plošší	7 913	120	12
1 137	plachý	plašší	4 454	23	0
1 410	hluchý	hlušší	2 889	14	0
1 892	prostoduchý	prostodušší	554	7	1
(5 123)	nejjednoduchý	nejjednodušší	196	0	1

(b) Adjektiva zakončená na ČK-í

Celá skupina vykazuje značnou variabilitu ve způsobech tvoření komparativu, ukazuje se jako vhodné rozčlenit ji do několika podskupin podle předposledního kmenového konsonantu.

(bI) s kmenovým zakončením -ZK (-žK)

Doloženo celkem 10 lemmat, jedno lemma má supletivní tvoření, u devíti se objevují dva různé typy redukce. Pro tvoření frekventovanějších komparativů je příznačná změna předposledního fonému Z > Ž, elize K a přípona -ší [ZK-í > Ž-ší],²⁵ zatímco méně četné komparativy jsou utvořeny změnou K > Č a příponou -í [ZK-í > ZČ-í]. Pozitiv *brzký* má supletivní komparativ *dřívější* (tvary *bržší/bržčí/bržčejší* nejsou vůbec doloženy; k hodnocení viz § 4.6.3.). U adjektiva *slizký* na opačné straně frekvenční škály se vyskytla varianta superlativu, která odpovídá produktivnímu tvoření u adjektiv s foneticky stejným, ovšem fonologicky odlišným zakončením SK-í.

²⁴ Tato podoba je navíc dále redukována: v komparativu se na morfematickém švu stýkají dvě stejné hlásky [š], které v řečovém procesu obvykle splývají v jednu; jde tedy v důsledku o změnu CH-í > Š-í.

²⁵ Stejně jako u třídy (a) je tato podoba dále redukována (viz pozn. 24). Pod vlivem asimilace znělosti se v komparativu na morfematickém švu stýkají dvě stejné hlásky [š] a ty se obvykle vyslovují dohromady; jde tedy v důsledku o změnu ZK-í > Š-í. – Pro úplnost analýzy dodávám: u stupňovaných tvarů s redukováním zakončením, které se tvoří od pozitivu s dlouhým vokálem v kořeni (*nížký, blízký, úzký*), dochází současně k další pravidelné redukci: zkrácení dlouhého vokálu (viz také dále *kratší, řidší*).

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
6	nízký	nižší	52 387	61 429	17 787
12	brzký	dřívější	7 325	17 934	20
17	blízký	bližší	46 554	13 633	46 416
20	těžký	těžší	99 989	11 235	10 176
31	úzký	užší	23 574	6 065	1 794
78	hezký	hezčí hezčejší	21 651	1 840 3	4 136 0
1 444	kluzký	kluzčí	2 095	13	3
3 957	mrzký	mrzčí	310	1	0
4 194	vazký	vazčí	108	1	0
4 707	slizký	slizčí slizštější	581	0 0	4 3

(bII) s kmenovým zakončením -TK

Doloženo celkem 5 lemmat, která vykazují tři různé způsoby tvoření. Jsou zde stejné způsoby tvoření jako u předchozí skupiny, ale tvary jsou napříč lemmaty mnohem dostupnější: čím vyšší frekvence komparativu, tím vyšší pravděpodobnost, že se komparativ bude tvořit elizí K a příponou -ší [TK-í > T-ší], s klesající frekvencí se komparativ tvoří spíše změnou kmenové hlásky K > Č a příponou -í [TK-í > TČ-í]; objevuje se i varianta s neredukovanou příponou -EJŠÍ [TK-í > TČ-EJŠÍ].

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
21	krátký	kratší	63 708	11 216	4 121
779	krotký	krotší krotčí krotčejší	874	45 3 2	6 0 0
1 016	vratký	vratší	1 411	28	8
1 507	plytký	plytčí plytčejší plytší	443	8 3 1	0 0 0
1 725	břítký	břitčí břitší	574	7 2	0 0

(bIII) s kmenovým zakončením na -DK

Doložena celkem 4 lemmata. Všechna mají komparativ tvořený elizí K a příponou -ší [DK-í > D-ší], omezeně se vyskytuje komparativ se změnou kmenové hlásky K > Č a s příponou -í [DK-í > DČ-í].²⁶ Tvoření neredukovanou příponou -EJŠÍ má jediný doklad.

²⁶ V této souvislosti stojí za pozornost frekvenční a strukturální rozdíly u dubletních stupňovaných tvarů se zakončením Č-í mezi třídami (bII) a (bIII). Na straně jedné v (bIII) varianta DČ-í nepřevyšuje u žádného z lemmat (*(nej-)prudčí, řidčí, sladčí, hladčí*) 2% podíl na počtu všech tvarů daného stupně, na straně druhé v (bII) se v lemmatech (s výrazně méně frekventovanými komparativy) prolínají dva způsoby redukováného tvoření; srov. také nízkou frekvenci tvarů (*(nej-)čistčí, hustčí, prostčí*) v třídě (d).

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
146	prudký	prudší prudčí prudčejší	20 010	660 10 1	478 2 0
166	řidký	řidší řidčí	3 422	543 10	15 0
208	sladký	sladší sladčí	13 414	396 8	335 7
349	hladký	hladší hladčí	12 438	201 2	58 0

(bIV) s kmenovým zakončením na -CK, -SK

Komparativy od rozsáhlých skupin adjektiv se zakončením na -CK-í (427 lemmat), resp. -SK-í (253 lemmat) se tvoří zcela pravidelně změnou posledních dvou kmenových hlásek CK > ČŤ, resp. SK > ŠŤ a produktivní příponou -EJŠÍ [CK-í > ČŤ-EJŠÍ; SK-í > ŠŤ-EJŠÍ]. V tabulce jsou obě skupiny zastoupeny lemmatem s nejvyšší *f* (KOMP).

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
122	optimistický	optimističtější	6 679	928	316
222	lidský	lidštější	118 502	369	75

(bV) s kmenovým zakončením na ostatní konsonanty: CK

Doloženo celkem 12 lemmat. Ačkoliv se jedná o lemmata s různorodými předposledními kmenovými konsonanty, nejedná se úplně o „zbytkovou“ kategorii. Lemmata vykazují měnící se vlastnosti v závislosti na frekvenci: (i) komparativ nejfrekventovanějšího adjektiva se tvoří supletivně; (ii) u dalších adjektiv se komparativ tvoří redukovaně změnou K > Č a příponou -í [CK-í > ČČ-í]; (iii) tvoření produktivní příponou -EJŠÍ je doloženo jen u méně frekventovaných adjektiv, a to okrajově (s výjimkou adj. *hořký*); (iv) nejméně čtený komparativ *horčejší* má doloženou jen neredukovanou příponu.

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
1	velký	větší	559 170	180 896	238 780
30	lehký	lehčí	34 983	6 296	745
102	měkký ²⁷	měkčí	12 786	1 204	93
140	tenký	tenčí tenší	10 858	719 13	181 8
216	křehký	křehčí	8 807	383	75
220	vlhký	vlhčí vlhčejší	7 900	378 2	36 0

²⁷ Lemma *měkký* zařazují do této třídy podle hlediska výsledného stavu i přesto, že současný fonemtický proces pozitiv > komparativ odpovídá změně K > KČ [MŇEK-í > MŇEKČ-í]. Žádné jiné adjektivum tyto znaky nevykazuje. Viz také pozn. 24 a 25.

406	mělký	mělčí mělčejší	2 783	161 1	8 1
523	trpký	trpčí trpčejší	2 858	106 2	60 0
555	křepký	křepčí křepčejší	63	96 1	0 0
942	hebký	hebčí hebčejší	1 186	30 4	10 0
1 150	hořký	hořčí hořčejší	6 560	16 6	6 6
2 844	horký	horčejší	27 093	2	1

(c) Adjektiva zakončená na VK-í

(cI) s kmenovým zakončením na OK

Doložena 4 lemmata. U tří vysoce frekventovaných komparativů se komparativ tvoří elizí skupiny OK, případným měkčením ($s > š$) a příponou -ší (SOK-í > š-ší, ROK-í > R-ší, BOK-í > B-ší), s výjimkou dubletního tvaru *hloubější*, který se patrně tvoří na základě komparativu adverbia. Nejméně frekventovaný komparativ se tvoří pouze změnou $k > č$ a neredukovanou příponou -EJší.

# KOMP	tvar POZ	tvar KOMP	$f(POZ)$	$f(KOMP)$	$f(SUP)$
2	vysoký	vyšší	197 525	122 063	108 259
10	široký	širší	46 471	21 022	4 271
24	hluboký	hlubší hloubější	35 459	9 062 5	3 200 4
217	divoký	divočejší	23 509	382	538

(cII) zájmenná adjektiva vyjadřující srovnání: *jinaký (jiný), onaký*

Doložena 2 lemmata. Adjektiva *jiný, jinaký, onaký* vyjadřují sama o sobě srovnání již v pozitivu, přesto formálně tvoří morfologické komparativní tvary (k tomu viz Haller – Ertl, 1925). Pro *jiný/jinaký* jsou doloženy všechny kombinatorické možnosti tvoření.

# KOMP	tvar POZ	tvar KOMP	$f(POZ)$	$f(KOMP)$	$f(SUP)$
173	jiný (jinaký)	jinačí jinší jinačejší jinější	570 932 (36)	419 92 16 1	0 0 0 0
464	onaký	onačejší onačí	203	134 3	1 0

(d) Adjektiva zakončená na ST-í

Doloženo celkem 11 lemmat (plus jedno záporné). Rozlišují strukturálně dvě podskupiny. (i) Adjektiva, která se stupňují změnou kmenového $T > ř$ a neredukovanou příponou

-EJŠÍ [ST-Í > SŤ-EJŠÍ]. (ii) Adjektiva, u nichž se vedle neredukovaného způsobu komparativ tvoří pomocí přípony -ŠÍ [ST-Í > ST-ŠÍ]. (Velmi omezenou variantou je tvoření ST-ČÍ.)

Redukovaný způsob tvoření u těchto lemmat výrazně převažuje nad neredukovaným. Přesto skoro 10% podíl neredukované varianty na celkovém dubletním tvoření komparativu je u adjektiv *čistý* a *hustý* relativně vysoký v porovnání se všemi ostatními dubletními neredukovanými komparativy v korpusu. Relativní frekvence neredukovaného komparativu pak stoupá až u těch nejméně častých (*sprostý*, *pustý*), což odpovídá očekávání na základě hypotézy (3). K těmto málo frekventovaným tvarům lze přiřadit i tvary záporného adjektiva *nečistý*, které lze ve vztahu k adjektivu *čistý* považovat za autonomní.

# KOMP	tvar POZ	tvar KOMP	f (POZ)	f (KOMP) % nered.	f (SUP)
40	častý	častější	18 503	4 213	9 612
65	čistý	čistší čistější čistčí	58 928	2 006 183 (8,35 %) 4	1 277 33 4
107	jistý	jistější	97 922	1 169	567
125	hustý	hustší hustější hustčí	12 746	776 86 (9,98 %) 0	361 28 2
186	prostý	prostší prostější prostčí	20 332	462 15 (3,14 %) 1	489 3 0
195	tlustý	tlustší tlustější tlustčí	12 304	430 12 (2,71 %) 1	156 1 0
680	sebejistý	sebejistější sebejistší	1 124	63 1	2 0
750	sprostý	sprostší sprostější sprostčí	3 304	43 10 (18,87 %) 0	116 4 1
1 116	pustý	pustší pustější	3 929	15 8 (34,78 %)	34 6
(1 974)	nečistý	nečistější nečistší	2 355	3 (50,00 %) 3	3 0
4 114	zlatistý	zlatistější	159	1	0
4 891	naprostý	naprostější	23 989	0	1

(e) Adjektiva zakončená na B-í

Doloženo celkem 15 lemmat. Podobně jako v (d) rozlišují dvě podskupiny. (i) Adjektiva odvozená ze substantiv a nominálních/předložkových frází, která se stupňují zcela pravidelně nezávisle na frekvenci vložení J ke kmeni a produktivní příponou -EJŠÍ [B-Í > BJ-EJŠÍ] (seřazeno sestupně podle frekvence: *dlouhodobý*, *krátkodobý*, *novodobý*, *soudobý*, *děledobý*, *bezzubý*, *velkohubý*). (ii) U ostatních adjektiv se tvoří kom-

parativ redukovanou příponou -ší [B-í > B-ší] a/nebo produktivním způsobem jako u první podskupiny. Relativní frekvence pravidelných komparativů na celkovém počtu komparativů u lemmatu stoupá s klesající četností komparativů, což odpovídá hypotéze (3). V tabulce nejsou uvedena méně frekventovaná lemmata (*libý, škodolibý, samolibý*), která se stupňují pravidelně.

# KOMP	tvar POZ	tvar KOMP	f (POZ)	f (KOMP) % nered.	f (SUP)
16	slabý	slabší slabější	29 502	13 968 2 (0,01 %)	4 006 1
141	hrubý	hrubší hrubější	42 526	703 28 (3,83 %)	569 13
294	blbý	blbější blbší	4 766	255 (98,84 %) 3	45 0
1 288	chabý	chabější	2 500	17 (100 %)	2
1 372	bohulibý	bohulibější	723	15 (100 %)	1

(f) Adjektiva zakončená na H-í

Doloženo celkem 6 lemmat. Tři různé způsoby tvoření jsou z větší části vysvětlitelné na základě frekvence. Nejfrekventovanější komparativ *dlouhý > delší* se tvoří supletivně (včetně negovaného *nedelší*); další dva komparativy jsou utvořeny se změnou H > Ž redukovanou příponou -ší [H-í > Ž-ší],²⁸ méně četné komparativy jsou utvořeny pravidelnou příponou -ejší. Z této řady se vymyká komparativ *nažší*, který lze jeho distribucí přirovnat ke komparativu *krasší*, v současných textech je nedoložen.

# KOMP	tvar POZ	tvar KOMP	f (POZ)	f (KOMP)	f (SUP)
8	dlouhý	delší nedelší	131 874	41 129 1	6 334
15	drahý	dražší	33 048	14 313	9 099
164	tuhý	tužší	9 052	557	152
668	ubohý	ubožejší	5 949	66	215
2 533	nahý	nažší	10 856	3	1
3 439	bezbřehý	bezbřežejší	1 243	1	1

(g) Adjektiva zakončená na ČR-í

Doloženo celkem 12 lemmat. Vyjma nejfrekventovanějšího komparativu *dobrý > lepší* se všechny komparativy od adjektiv, které jsou v kmeni zakončeny na ČR, tvoří se změnou v poslední kmenové hláске R > Ř produktivní příponou -ejší [ČR > ČŘ-ejší]. Všechna lemmata podle klesající frekvence: *ostrý, chytrý, moudrý, pestrý, štědrý, bystrý, příkrý, modrý, kyprý, mokrý, bodrý*. (Pro variantu bez měkčení existuje jediný doklad: *příkrejší*, viz také dále odst. (h).)

²⁸ Stejně jako u tříd (a) a (b) je podoba komparativu dále redukována (viz pozn. 24, 25). Jde tedy v důsledku o změnu H-í > š-í.

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
3	dobrý	lepší lepčí dobřejší lepčejší lepšejší	254 145	115 872 66 9 5 3	166 102 36 7 3 1
602	příkrý	příkřejší příkřejší	2 351	80 1	36 0

(h) Adjektiva zakončená na VR-í

Doloženo celkem 5 lemmat (plus jedno záporné). Ve skupině se vyskytuje jeden vysoce frekventovaný redukovaný komparativ *starší*, další komparativy se tvoří stejně jako u adjektiv na CR-í, ale s tím rozdílem, že u některých lemmat nedochází dubletně ke změně v kmeni R > Ř.

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
5	starý	starší	213 411	76 270	25 958
340	přímočarý	přímočařejší přímočarejší	2 016	211 2	1 0
958	ponurý	ponuřejší ponurejší	2 158	24 9	10 1
1 280	bujarý	bujarejší bujarejší	1 276	14 3	5 3
2 406	jarý	jařejší	144	4	0
(4 562)	nepřímočarý	nepřímočařejší	8	1	0

(i) Adjektiva zakončená na L-í

Všechny komparativy od adjektiv zakončených v kmeni na -L (celkem 347 lemmat) se tvoří pravidelně pomocí přípony -EJŠÍ, s jedinou výjimkou, a to supletivního tvoření komparativu *malý* > *menší* (včetně *nemenší*). V tabulce je pro doplnění i počet pravidelných stupňovaných dokladů od adjektiva *zlý*, pro něž se v gramatikách (již např. v Gebauerovi 1895, s. 61) obvykle uvádí komparativ *horší* (jeho frekvenční charakteristiky viz tabulku pod odst. (k)).

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
4	malý	menší nemenší	274 988	78 920 324	23 967
1 171	zlý	zlejší	26 602	21	19

(j) Adjektiva zakončená na D-í

V rámci této třídy adjektiv vymezují podtřídu s vokálem E na předposledním místě v kmeni, v níž se způsob vyjádření stupňování od zbytku adjektiv z této třídy významně odlišuje.

(jI) s kmenovým zakončením -ED

Celkem 6 dokladů. Vyjma jediného dokladu se všechny hlavní tvary komparativu tvoří příponou -ší [D-í > D-ší]. U adjektiv *škaredý* a *rudohnědý* je doložen produktivní způsob tvoření se změnou poslední kmenové hlásky D > Ď příponou -EJŠÍ [D-í > Ď-EJŠÍ].

# KOMP	tvar POZ	tvar KOMP	f (POZ)	f (KOMP)	f (SUP)
355	bledý	bledší	5 459	198	4
579	snědý	snědší	2 181	89	0
1 502	hnědý	hnědší	13 666	12	0
1 531	škaredý	škaredší škaredější	602	6 5	6 8
1 704	šedý	šedší	12 559	9	1
4 125	rudohnědý	rudohnědější	152	1	0

(jII) s jiným kmenovým zakončením než -ED

Celkem 12 dokladů. Frekvencí a formálním značením se zde jednoznačně vydělují dvě podskupiny. (i) Tři nejfrekventovanější komparativy se tvoří příponou -ší [D-í > D-ší]. (ii) Všechny zbylé komparativy mají výrazně nižší frekvenci a s jedinou výjimkou produktivní neredukované tvoření. Seznam lemmat podle klesající frekvence: *různorodý*, *hrdý*, *rudý* (avšak 1 doklad *nejrudší*), *čínorodý*, *sourodý*, *stejnorodý*, *prvořadý*, *druhořadý*, *cizorodý*. Srov. s třídou (bI).

# KOMP	tvar POZ	tvar KOMP	f (POZ)	f (KOMP)	f (SUP)
7	mladý	mladší	198 295	44 052	12 780
32	tvrdý	tvrdší	47 414	6 027	2 621
43	chudý	chudší	20 136	3 848	3 061

(k) Adjektiva zakončená na N-í

Komparativ se od této vůbec nejpočetnější skupiny stupňovaných adjektiv (2 567 lemmat) tvoří zcela pravidelně změnou kmenové hlásky N > ň a příponou -EJŠÍ [N-í > ň-EJŠÍ]. Tomu se vymykají pouze čtyři případy: supletivní komparativ od *špatný*, tvoření komparativu od *snadný*, básnický tvar *krasší* a výše uvedené *jiný* (cII).

U lemmatu *snadný* lze jako u jediného vysoce frekventovaného komparativu pozorovat skutečně dvojí užívání různých tvarů: vedle neredukovaného komparativu je zde častější tvoření se změnou v kmeni D > Z, elizí N a příponou -ší [DN-í > Z-ší] (srov. adv. *snáze*). Neredukovaná přípona zde vznikla až druhotně ve staré češtině vzácným diachronním procesem. Zatímco u řady komparativů v té době nově vznikaly redukované podoby (Marvan, 2008, s. 174), u omezeného počtu adjektiv (se zakončením kmene na DN, vedle *snadný* např. *zadní*, *přední*, *poslední*; dále také *krásný*) se naopak ke starším redukovaným vytvořily analogické neredukované varianty. U většiny těchto lemmat redukované varianty dodnes zanikly nebo získaly zvláštní význam; viz *zadní* v odst. (I). U lemmatu *snadný* redukovaná varianta přetrvala, patrně se zde diachronně projevil konzervační efekt. Oproti tomu záporné adjektivum *nesnadný* podlehl regularizaci –

jeho stupňování je doloženo pouze produktivní příponou (vůči adjektivu *snadný* se tak jeví jako zcela autonomní).

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
9	špatný (zlý?)	horší špatnější	85 711	36 748 5	29 925 2
27	snadný	snazší snažší snadnější	22 107	4 497 288 2 798	517 27 471
68	krásný	krásnější krasší	68 726	2 084 1	12 015 5
(559)	nesnadný	nesnadnější	2 289	95	21

(l) Adjektiva zakončená na ň-í

Komparativ od adjektiv na ň-í (663 lemmat) se tvoří stejným způsobem jako u adjektiv zakončených na N-í s tím rozdílem, že kmenová hláska zůstává nezměněna [Ň-í > Ň-EJŠÍ].²⁹ Z této skupiny se vydělují lemmata *pozdní*, *horní* a *dolní*,³⁰ u nichž se komparativ bez hlásky ň tvoří na základě tvaru adverbia, tedy vlastně supletivně (srov. § 4.6.3., *dřívější*). Tvary *hořejší* a *dolejší* však fungují spíše jako dublety k pozitivu (srov. *tamní*, *tamější*). Stupňování u *zadní* se tvoří podobně jako u výše uvedeného *snadný* (ovšem adv. *záze* „více vzadu“ není součástí dnešního úzu). Pro porovnání uvádím i stupňování lemmatu *přední*.

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
11	pozdní	pozdější pozdnější	11 443	18 974 71	72 27
181	přední	přednější	47 117	505	511
1 657	zadní	zazší zažší zadnější	28 811	2 0 7	1 444 38 26
(298) 3 260	horní	hořejší hornější	48 637	(256) 1	118 18
(494) 2 964	dolní	dolejší dolnější	34 250	(121) 2	9 0

(m) Adjektiva zakončená na VT-í

Komparativ se u adjektiv na VT-í (320 lemmat) tvoří nezávisle na frekvenci zcela pravidelně, a to se změnou kmenové hlásky T > Ť příponou -EJŠÍ [T-í > Ť-EJŠÍ]. Třemi izolovanými případy jsou adjektiva, u nichž se tvoří komparativ příponou -ŠÍ [VT-í > VT-ŠÍ].

²⁹ Z této skutečnosti také vyplývá, že některá (často sémanticky velmi podobná) adjektiva mají homonymní stupňované tvary, viz např. *čelní* – *čelný*. V rámci provedené analýzy se tak lemmata mohou překrývat.

³⁰ Lemma *pozdní* je také redukované (viz výše § 2.1.). Na rozdíl od ostatních tvarů však nemá komparativ redukovanou příponu, nýbrž pouze kmen. Vzhledem k nejednoznačné sémantice mezi redukované nezařazují komparativy *hořejší* a *dolejší*.

U vysoce frekventovaného komparativu od *bohatý* se jedná o základní tvoření, zatímco u adjektiv *krutý* a *žlutý* jde o vedlejší varianty. Motivaci pro toto tvoření lze snad hledat u sémanticky a hláskově podobných komparativů (*krotší*, resp. *hnědší*, *šedší*).

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
34	bohatý	bohatší bohatější	43 052	5 117 64	6 179 3
211	krutý	krutější krutší	12 587	400 1	603 3
957	žlutý	žlutější žlutší	28 169	19 14	3 1

(n) Adjektiva zakončená na v-í

Třída adjektiv na v-í opět zahrnuje vysoký počet adjektiv (698 lemmat), která jsou stupňována pravidelně vložením J ke kmeni a produktivní příponou -EJŠÍ [v-í > vJ-EJŠÍ]. Jediným izolovaným redukováným lemmatem je *tmavý*.

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
91	tmavý	tmavší tmavější	15 065	1 452 18	100 7

(o) Adjektiva zakončená na ouc-í, íc-í

Zvláštní třídu tvoří adjektiva odvozená ze slovesných tvarů přechodníků přítomných (celkem doloženo 20 lemmat).³¹ Stupňují se pravidelně vložením ň ke kmeni a příponou -EJŠÍ, patrně na základě tvarů adverbii. Lemmata s nejčetnějšími komparativy jsou uvedena v tabulce. V některých případech jsou doloženy dubletní tvary pozitivu se zakončením n-í (viz zejm. *vroucný*, *horoucný*, *kajícný*, *nelpřejícný*, *nemohoucný*). Srov. Stich (1969).

# KOMP	tvar POZ	tvar KOMP	<i>f</i> (POZ)	<i>f</i> (KOMP)	<i>f</i> (SUP)
1 008	žádoucí	žádoucnější	7 214	28	66
1 028	vroucí vroucný	vroucnější	1 735 113	27	69
1 700	vědoucí	vědoucnější	463	9	2
1 737	horoucí horoucný	horoucnější	633 6	8	14

(p) Adjektiva s ostatními zakončeními

Zbývá adjektiva se zakončeními, která dosud nebyla zmíněna, se stupňují pravidelně s příponou -EJŠÍ. Pro úplnost uvádím jejich výčet: m-í, př. *známý* [m-í > mŇ-EJŠÍ] (20 lemmat); p-í, př. *hloupý* [p-í > pJ-EJŠÍ] (8 lemmat); s-í, př. *kosý* [s-í > s-EJŠÍ] (6 lemmat); z-í, př. *drzý*, *cizí* [z-í > z-EJŠÍ] (3 lemmata); ž-í, př. *svěží* [ž-í > ž-EJŠÍ] (1 lemma).

³¹ Za upozornění na tuto skupinu adjektiv děkuji jednomu z anonymních recenzentů tohoto textu.

4.6.2. Interpretace dat: redukční a konzervační efekt v jednotlivých třídách

Úplný seznam tvarů s uvedením jejich frekvence v § 4.6.1. poukázal na zřejmou (ovšem predikovanou) asymetrii v systému českého stupňování. (i) Typově nejčastější produktivní vzorec (k) se zakončením -N-í systematicky nepodléhá celoskupinovému redukčnímu efektu (1). (ii) V češtině na druhé straně existují malé uzavřené třídy adjektiv (CH-í, ED-í) se zcela pravidelným (v rámci uzavřené třídy „produktivním“) tvořením komparativu pomocí redukované přípony (-ší); neredukovaný typ stupňování nepřipouští ani adjektiva se zakončením DK-í. (iii) Mezi těmito dvěma strukturně vymezenými extrémy existuje řada tříd adjektiv, v jejichž rámci je způsob tvoření komparativu (redukovaný oproti neredukovanému) do velké míry vysvětlitelný frekvencí. Vzhledem k nízkému počtu dokladů u některých stupňovaných tvarů však musím s frekvenčními hodnotami z § 4.6.1. nakládat obezřetně, nekladu proto důraz na některé sice na první pohled nápadné, avšak statisticky nesignifikantní frekvenční vztahy.

Ukázkou vlivu frekvence na strukturu stupňování jsou např. rozdíly v třídách (bI) *nižší vs. slizčí*, (bII) *kratší vs. břitčí*, (cI) *vyšší vs. divočejší*, (f) *delší vs. ubožejší*, (jII) *mladší vs. hrdější*. Ve světle frekvenční analýzy není rovněž náhodné, že právě u nejfrekventovanějších adjektiv z tříd (g), (i), (k), kde jinak nedochází k systematické redukci, se ustálily svébytné supletivní krátké tvary komparativů: *dobrý > lepší, malý > menší, špatný > horší* (viz § 4.3.).

Podle hypotézy redukčního efektu vnitroskupinového (3) o vztahu redukovaných tvarů k neredukovaným v rámci jednoho lemmatu jsem očekával, že u lemmat s vyšší absolutní frekvencí všech stupňovaných tvarů bude poměr redukovaných tvarů vůči neredukovaným vyšší než u lemmat s nižší absolutní frekvencí. Tento předpoklad se zřetelně prokázal v rámci třídy (e) (B-í > B-ší/BJ-EJší), částečně v třídách (bV) a (d). Role tohoto efektu je však v rámci systému českého stupňování spíše okrajová. Co se totiž ukázalo jako závažné, a to i vzhledem k otázkám *ja z y k o v é n o r m y*, u naprosté většiny lemmat s dubletním stupňováním (viz především třídy (bII), (bV) a (d)) jsou vedlejší neredukované tvary v porovnání s hlavními redukovanými tvary zcela okrajové. Data neprokazují, že by v českém stupňování (alespoň u některých tříd) postupovala regularizace. Konzervační efekt o zpravidelnování (4) se v češtině projevuje jen do té míry, že neredukovanému tvoření někdy podléhají v rámci tříd adjektiv se stejným zakončením jen ty nejméně frekventované tvary komparativu. Dále: v korpusu jsou doloženy jen velmi vzácně neredukované tvary známé z gramatik (*hebčejší, křepčejší, vlnčejší*). Jiné existují v korpusu pouze v podobě dokladů v rámci učebních materiálů (*měččejší*) nebo nejsou doloženy vůbec (*křehčejší*).³² Jediným vysoce frekventovaným lemmatem se skutečně dubletním stupňováním je tak adjektivum *snadný* (viz odlišný způsob vzniku neredukovaného komparativu), zaznamenaníhodný podíl vedlejších neredukovaných tvarů komparativu mají dále už jen lemmata *čistý* a *hustý*.

Místo o regularizaci tvarů komparativu je tedy spíše na místě hovořit o historicky zakonzervovaných způsobech stupňování v jednotlivých třídách, a ty jsou do značné míry vysvětlitelné frekvencí tvarů komparativu a působením redukčního efektu.

³² V této souvislosti upozorňuji, že při výzkumu jsem narazil i na několik nesrovnalostí v internetové jazykové příručce Ústavu pro jazyk český AV ČR (viz <<http://prirucka.ujc.cas.cz/>>, např. jsou zde uvedeny jako jediné varianty komparativu v korpusu nedoložené tvary *snědější, brzčejší*, cit. 10. 6. 2010).

4.6.3. Supletivismus prostupující třídami adjektiv

Supletivismus jako nejsilnější důsledek redukčního efektu a tvarové autonomie se v češtině projevuje velmi omezeně. V nejsilnější podobě (střídání tvarů původně nepřibuzných slov v jednom paradigmatu) jej lze pozorovat jen u lemmat *velký, dobrý, malý, špatný* a *brzký*. U lemmat *dlouhý, pozdní* je kmen komparativu příbuzný s kmenem pozitivu. Možnost vysvětlení takto malého počtu supletivních tvarů jsem nastínil již v § 4.6.2.: supletivní tvary se ujaly u frekventovaných komparativů těch tříd, v nichž se nevyvinula systematická redukce zkrácenou příponou.

Jiným případem je lemma *brzký*. Vysvětlení pro supletivismus *brzký – dřívější* je třeba hledat ve vztazích mezi adverbii, z nichž jsou tvary tohoto lemmatu odvozeny (Marvan, 2008, s. 81). Dochází zde ke stejnému posunu jako u dvojice *pozdní – pozdější*, posun je však formálně komplikován a determinován supletivismem v rámci dvojice *brzy – dříve*; ta funguje na místo dvou teoreticky možných synonymních řad tvarů *brzy – [?]brzeji (brzký – [?]brzejší)* a *dříve – [?]dřívěji ([?]dřevní – dřívější)*. Srov. v *dřevních dobách*. Možný tvar *brzcí/brzčejší* nemohu v dnešní češtině vyloučit, vzhledem k existenci konkurenčního lemmatu *časný* jej však považuji za velmi nepravděpodobný. Srov. k tomu § 4.1. a další vztahy v kategorii (I) (*hořejší, dolejší*).

Náznačky supletivismu pozorují také u dalších lemmat. Zaměřím se zde na dvě adjektiva hodnotící krásu: *hezký* a *pěkný*.

Tab. 3: Porovnání stupňování *hezký – pěkný*

# KOMP	tvar POZ	tvar KOMP	f (POZ)	f (KOMP)	f (SUP)
78	hezký	hezčí	21 651	1 840 (6,66 %)	4 136 (14,97 %)
510	pěkný	pěknější	28 324	114 (0,40 %)	284 (0,99 %)
68	krásný	krásnější	68 726	2 084 (2,52 %)	12 015 (14,51 %)
616	nádherný	nádhernější	19 026	78 (0,40 %)	206 (1,07 %)

(Procenta za frekvencí KOMP a SUP v tabulce udávají relativní frekvenci tvaru vůči všem tvarům lemmatu. Pro srovnání jsou uvedena i další dvě sémanticky spřízněná, ovšem významově výrazněji odstíněná lemmata *krásný* a *nádherný*.)

Tato adjektiva jsou v pozitivu podobně frekventovaná a lze je považovat víceméně za synonyma bez silnějšího významového odstínu. To však neplatí pro jejich stupňované tvary. Stupňování je vyjadřováno v drtivé většině pomocí redukovaných tvarů adjektiva *hezký*. Považuji to za další projev redukčního efektu v češtině: výběr kratšího tvaru komparativu oproti synonymnímu tvaru delšímu.

5. Shrnutí a závěry

V příspěvku jsem ukázal, jak podrobné prozkoumání frekvence jazykových jednotek může do značné míry přispět k pochopení a vysvětlení jejich morfematické stavby, vztahů mezi jednotkami navzájem, univerzálních vývojových tendencí i možných kognitivních principů, které tyto vztahy utvářejí. Zjištěná frekvence jazykových tvarů – rozdělení jednotek na méně a více časté – nebyla v tomto příspěvku cílem popisu (ne-

měla ukázat, jak jazyk skutečně vypadá), nýbrž ukázala se být vhodnou proměnnou, která indikuje a vysvětluje, proč jazyk vypadá právě tak, jak vypadá.

Zaměřil jsem se zejména na to, jak se v českých datech projevují univerzální frekvenční efekty. Jejich teoretické zakotvení v rámci tzv. přístupů založených na užívání (funkcionalistického typologického výzkumu) jsem představil v úvodní části článku. Na tomto základě jsem pak provedl korpusový průzkum a zformuloval jednotlivé jazykově specifické hypotézy pro materiál českého jazyka.

Po představení základních deskriptivních (strukturních a frekvenčních) charakteristik tvarů českého stupňování a při jejich zpřesňování jsem dospěl k následujícím zjištěním o vztahu frekvence a jejich morfemické struktury:

- Ukázala se výrazná role absolutní frekvence a tzv. redukčního efektu (§ 4.1.). Redukované tvary komparativu patří ke komparativům s nejčastějším výskytem, mají celkově vysokou tokenovou frekvenci; jejich typová frekvence je oproti tomu velmi nízká. Patrný je v češtině i efekt autonomie: supletivismus jako krajní důsledek tohoto efektu je charakteristický pro adjektiva s nejméně frekventovanějšími komparativy.
- Prozkoumal jsem (v § 4.2.) otázku, jak morfosyntaktická asymetrie mezi tvary českého stupňování (superlativní tvary jsou delší než tvary komparativní, mají pravidelné tvoření a nepodléhají frekvenčním efektům) odpovídá predikované univerzální tendenci v Haspelmathově pojetí. Česká data tuto tendenci podporují: superlativní tvary jsou z hlediska relativní frekvence méně frekventované než tvary komparativní, což koreluje s očekáváním.
- Pomocí efektu autonomie a stupňů produktivity podle pojetí J. L. Bybeeové jsem ukázal možnost, jak pojímat kategoriální vztahy mezi českými stupňovanými tvary (§ 4.3., § 4.4., § 4.5.), především jsem ukázal roli produktivní přípony *-ejší*. Tímto pojetím jsem se mj. vyhnul otázce přiřazení stupňování ke gramatice, nebo lexikonu (která se tím stává nerelevantní).

Nakonec jsem se věnoval hlavnímu materiálovému těžišti této práce. V § 4.6.1. jsem uvedl a klasifikoval všechny doložené způsoby tvoření komparativu,³³ a to podle strukturních i frekvenčních hledisek. Na základě toho jsem formuloval tyto nejdůležitější závěry, které se týkají především redukčního a konzervačního efektu (podrobně jsou rozvedeny v § 4.6.2. a § 4.6.3.):

- Nejčastějším produktivním vzorcem jsou adjektiva zakončená na *-ný*, která systematicky nepodléhají celoskupinovému redukčnímu efektu. Dále existují malé uzavřené třídy, u nichž se komparativ tvoří zcela pravidelně redukovanou příponou. Pro řadu dalších tříd adjektiv platí, že mají redukovaný i neredukovaný způsob tvoření komparativu, který je vysvětlitelný frekvencí: adjektiva s komparativy o vyšší frekvenci mají tendenci je mít redukované, zatímco adjektiva se stejným zakončením, ale s méně frekventovanými komparativy mají pravidelné stupňování.
- Role redukčního efektu vnitroskupinového, podle něhož by u lemmat s dubletním stupňováním měl podíl neredukovaných komparativů stoupat s klesající tokenovou

³³ Přesněji řečeno: analyzoval jsem všechny způsoby tvoření komparativu, které se podařilo odhalit v korpusu SYN pomocí procedury popsané v § 2.

frekvencí komparativu, se ukázala jako okrajová. U naprosté většiny lemmat s dubletním stupňováním jsou totiž vedlejší neredukované tvary v porovnání s hlavními redukovanými tvary velmi málo frekventované. Ačkoliv jsou dubletní neredukované tvary z hlediska jazykové normy kodifikované, jsou jen minimálně využívány, někdy jsou dokonce nedoložené. S tím souvisejí i očekávání podle konzervačního efektu. V rámci tříd adjektiv s dubletními komparativy by mohlo docházet ke zpravidelňování u lemmat s nižší frekvencí komparativu. To se však nepotvrdilo. Jen v několika třídách platí, že neredukovanému tvoření podléhají v rámci třídy pouze ty nejméně frekventované tvary komparativu.

- Suppletivismus se napříč třídami projevuje jen velmi omezeně, supletivní tvary komparativu se ujaly především u těch zakončení, kde nedochází k systematickému redukovanému tvoření. Náznaky suppletivismu lze také pozorovat ve frekvenčních asymetriích mezi dvojicemi synonymních adjektiv.

Frekvence se tedy ukázala jako vhodný analytický nástroj. Podrobným zkoumáním frekvenčních efektů na materiálu z jazykových korpusů je možné v jazyce leccos ozřejmit a vysvětlit. Jak jsem však několikrát ukázal, konkrétní hodnoty ukazatelů nelze přeceňovat, vždy musejí být vzaty v úvahu i podmínky a souvislosti, za nichž byla frekvenční data získána. Do budoucna považuji za užitečné prozkoumat z této perspektivy i další oblasti české gramatiky.

ZDROJE

Český národní korpus – SYN2000. Praha: Ústav Českého národního korpusu FF UK, 2000. Dostupný online na adrese: <<http://www.korpus.cz>>.

Český národní korpus – SYN2005. Praha: Ústav Českého národního korpusu FF UK, 2005. Dostupný online na adrese: <<http://www.korpus.cz>>.

Český národní korpus – SYN2006PUB. Praha: Ústav Českého národního korpusu FF UK, 2006. Dostupný online na adrese: <<http://www.korpus.cz>>.

LITERATURA

BYBEE, J. L. (1985): *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.

BYBEE, J. L. (1995): Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, s. 425–455. Přetištěno in: J. L. Bybee (2007a, s. 167–193).

BYBEE, J. L. (2007a): *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

BYBEE, J. L. (2007b): Introduction. In: J. L. Bybee, *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press, s. 5–22.

CROFT, W. (1990): *Typology and Universals*. Cambridge: Cambridge University Press.

CUZZOLIN, P. – LEHMANN, Ch. (2004): Comparison and gradation. In: G. E. Booij – Ch. Lehmann – J. Mugdan (eds.), *Morphologie: Ein Internationales Handbuch zur Flexion und Wortbildung*, 2 / *Morphology: An International Handbook on Inflection and Word-Formation*, 2. Berlin: Mouton de Gruyter, s. 1212–1220.

ČVRČEK, V., a kol. (2010): *Mluvnice současné češtiny*, 1. Praha: Karolinum.

ČECH, R. (2005): Komunikace versus systém, nebo komunikace versus model? *Slovo a slovesnost*, 66, s. 176–179.

- ČECH, R. (2007): Language system – linguistics as an empirical science. *Sapostavitelno ezikoznanie*, 32, s. 42–49.
- DANEŠ, F. (1966): The relation of centre and periphery as a language universal. In: *Travaux linguistiques de Prague, 2: Les problèmes du centre et de la périphérie du système de la langue*. Praha: Academia, s. 9–21.
- GEBAUER, J. (1895): *Mluvnice česká pro školy střední a ústavy učitelské: Nauka o slově, I*. Praha: Tempský.
- GREENBERG, J. H. (ed.) (1966): *Universals of Language*. Cambridge: The MIT Press.
- HALLER, J. – ERTL, V. (1925): Jinší, jinačí. *Naše řeč*, 9, s. 289–297.
- HAIMAN, J. (1983): Iconic and economic motivation. *Language*, 59, s. 781–819.
- HASPELMATH, M. (2008a): Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics*, 19, s. 1–33.
- HASPELMATH, M. (2008b): Creating economical morphosyntactic patterns in language change. In: J. Good (ed.), *Linguistic Universals and Language Change*. Oxford: Oxford University Press, s. 185–214.
- HASPELMATH, M. (2008c): Reply to Croft and Haiman. *Cognitive Linguistics*, 19, s. 59–66.
- HOPPER, P. (1987): Emergent grammar. In: *Proceedings of the Annual Meeting of the Berkeley Linguistics Society, 13*. Berkeley: Berkeley Linguistics Society, s. 139–157.
- JAKOBSON, R. (1970): Hledání podstaty jazyka. In: É. Benveniste (ed.), *Dvanáct esejů o jazyce*. Praha: Mladá fronta, s. 29–45.
- KARLÍK, P. – HLADKÁ, Z. (2004): Kam s ním? (Problém stupňování adjektiv). In: P. Karlík – J. Pleskalová (eds.), *Život s morfémy: Sborník studií na počest Zdenky Rusínové*. Brno: Masarykova univerzita, s. 73–93.
- KARLÍK, P. – NEKULA, M. – RUSINOVÁ, Z. (eds.) (2003): *Příruční mluvnice češtiny*. Praha: Nakladatelství Lidové noviny.
- KOŘENSKÝ, J. (1987): K procesuálnímu modelování řečové činnosti. *Slovo a slovesnost*, 48, s. 177–189.
- KOŘENSKÝ, J. (1994): Morfologické kategorie a procesuálně-komunikační přístup k řeči a jazyku. *Slovo a slovesnost*, 55, s. 81–89.
- KOŘENSKÝ, J. (2003): Procesuální gramatika v kontextu současných tendencí lingvistického myšlení. *Slovo a slovesnost*, 64, s. 1–7.
- MARVAN, J. (2008): *České stupňování – slovanská podoba evropské univerzálie*. Praha: Karolinum.
- REJZEK, J. (2001): *Český etymologický slovník*. Voznice: Leda.
- ROSCH, E. (1977): Classification of real-world objects: origins and representations in cognition. In: P. N. Johnson-Laird – P. C. Wason (eds.), *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press, s. 212–222.
- STICH, A. (1969): Stupňování přídavného jména vzrušující. *Naše řeč*, 52, s. 62–63.
- TAYLOR, J. R. (1989): *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Oxford University Press.
- ÚČNK (2010): Korpus SYN2006PUB. Praha: Ústav Českého národního korpusu FF UK. Dostupné online na adrese: <<http://ucnk.ff.cuni.cz/syn2006pub.php>>. Cit. 20. 6. 2010.
- VYKYPĚL, B. (2009): *Empirical Functionalism and the Prague School* (= Travaux linguistiques de Brno, 6). München: LINCOM Europa.
- ZIPF, G. K. (1929): Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology*, 40, s. 1–95.
- ZIPF, G. K. (1949): *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge: Addison-Wesley.

Ústav lingvistiky a ugrofinistiky FF UK
 nám. Jana Palacha 2, 116 38 Praha 1
 <jan.krivan@ff.cuni.cz>