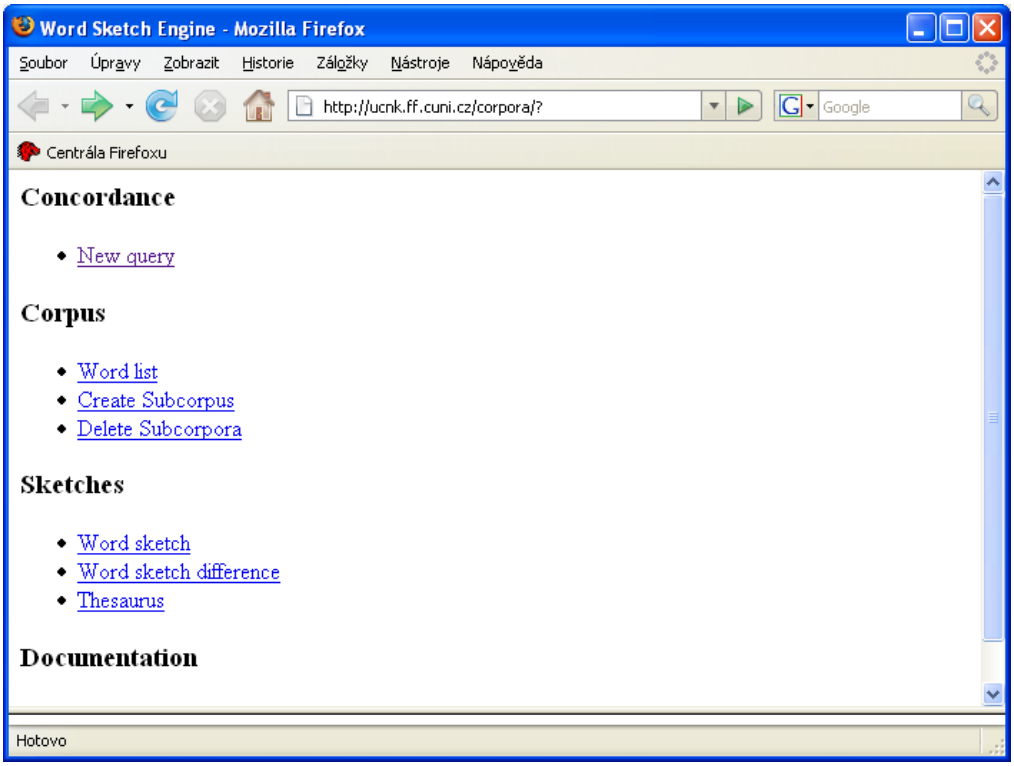
Začínáme s Bonitem 2
//Word Sketch Engine//

Východisko
Bonito 2 je internetový program, který lze použít na zpracování korpusu libovolného jazyka, je-li tento korpus označkován vhodným způsobem.
Bonito 2 má řadu funkcí, z nichž základní jsou:

konkordancer (velmi rychlý a vysoce funkční)
program Word Sketch (viz dále).

Více informací o programu Word Sketch naleznete na Kilgarriff et al 2004 in Proc EURALEX.

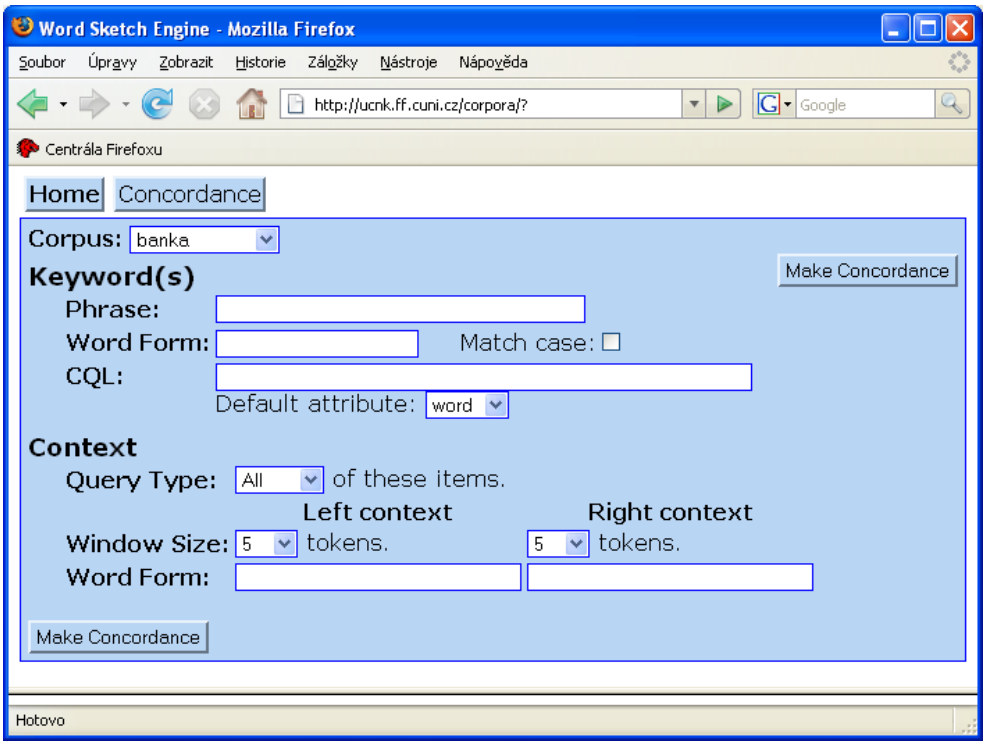
Domovská stránka
Domovskou stránku jste zobrazili, když jste se do programu přihlásili jako registrovaní uživatelé Českého národního korpusu. Jste tedy na stránce:



Concordance – New query

1.1. Konkordance – Nový dotaz

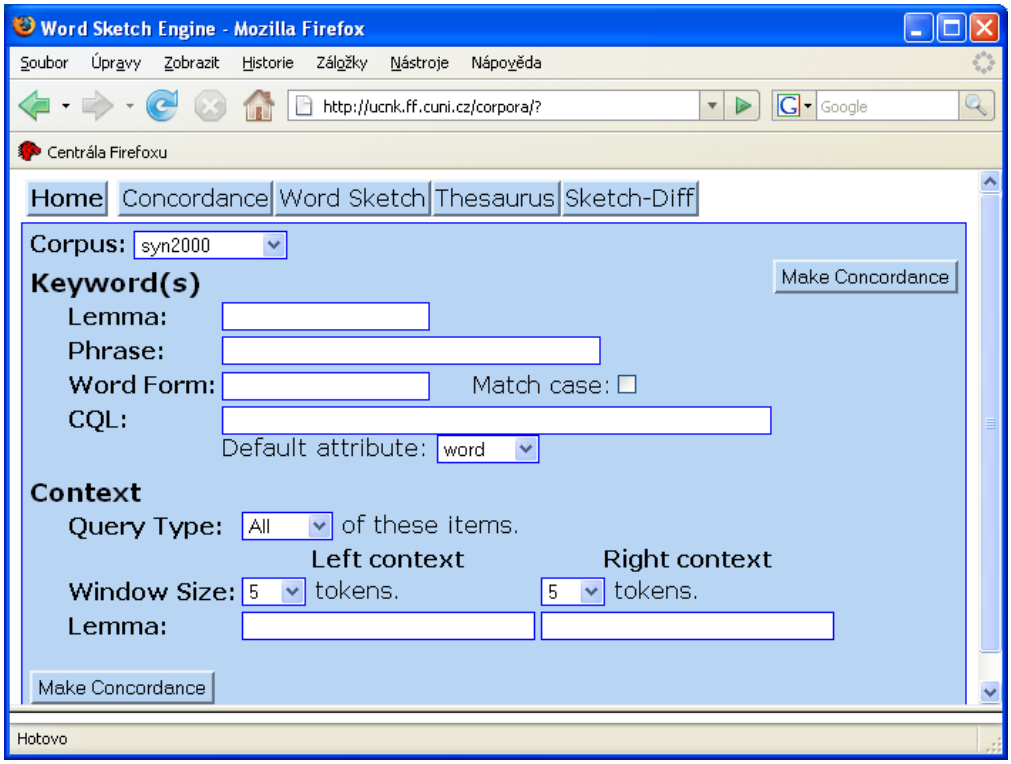
Klikněte na „Concordance - New Query“, dostanete se na stránku zadávacího formuláře.



Nad formulářem je tlačítko pro návrat na domovskou stránku „**Home**“.

Na formuláři je vedle vodícího ukazatele „**Corpus**“ rozevírací roletová nabídka korpusů, se kterými lze aktuálně pracovat. Vytvoříte-li si subkorpus, bude se vám v budoucnu zobrazovat v této nabídce také.

V této nabídce si vybereme korpus SYN2000. Tím se nám stránka změní na:



1.2. Hledané slovo / hledaná slova

Do sekce „**Keyword(s)“** budeme zadávat námi hledané slovo / hledaná slova.

Za vodící ukazatel „**Lemma**“ zadáváme dotaz, hledáme-li přes lemma. Za vodicí ukazatel „**Word**“ naopak zadáváme dotaz, hledáme-li jednu konkrétní formu, necháme-li přitom prázdné pole „Match case“, bude se tvar hledat jak s malými, tak s velkými písmeny (v Bonitu 1 hledání „lc“). Zaklikneme-li možnost „Match case“, bude se forma hledat přesně tak, jak jsme ji zadali, tedy např. jen se všemi písmeny malými.

Za vodící ukazatel „**Phrase**“ zadáváme dotaz, hledáme-li kombinace slov, v tomto poli jde však vždy o kombinace sousední. Lze vyhledávat jak na úrovni wordů (je základně nastavena za vodícím ukazatelem „Default attribute“, stejně jako je základně nastavena libovolná velikost písmen díky prázdnému poli „Match case“), tak na úrovni lemmat. Tuto si můžeme nastavit za vodícím ukazatelem „Default attribute“.

Za vodícím ukazatelem „**CQL**“ lze zadávat dotazy pomocí regulárního jazyka. Je to ta forma dotazu, která se nám zobrazila v dotazovacím řádku Bonita 1, vytvořili-li jsme si složitější dotaz pomocí „Grafického vytváření“. Nepamatujeme-li si pravidla vytváření dotazu v tomto jazyce, můžeme dotaz vytvořit v „Grafickém vytváření“ Bonita 1, dotaz odeslat do korpusu, takže se nám znovu zobrazí v dolním okně. Z tohoto okna pak lze kopírovat.

Za vodícím ukazatelem „**Default attribute**“ si můžeme pro jednoslovné i víceslovné dotazy vybrat, zda chceme hledat na úrovni formy („Word“), nebo na úrovni lemmatu („Lemma“).

1.3. Kontext

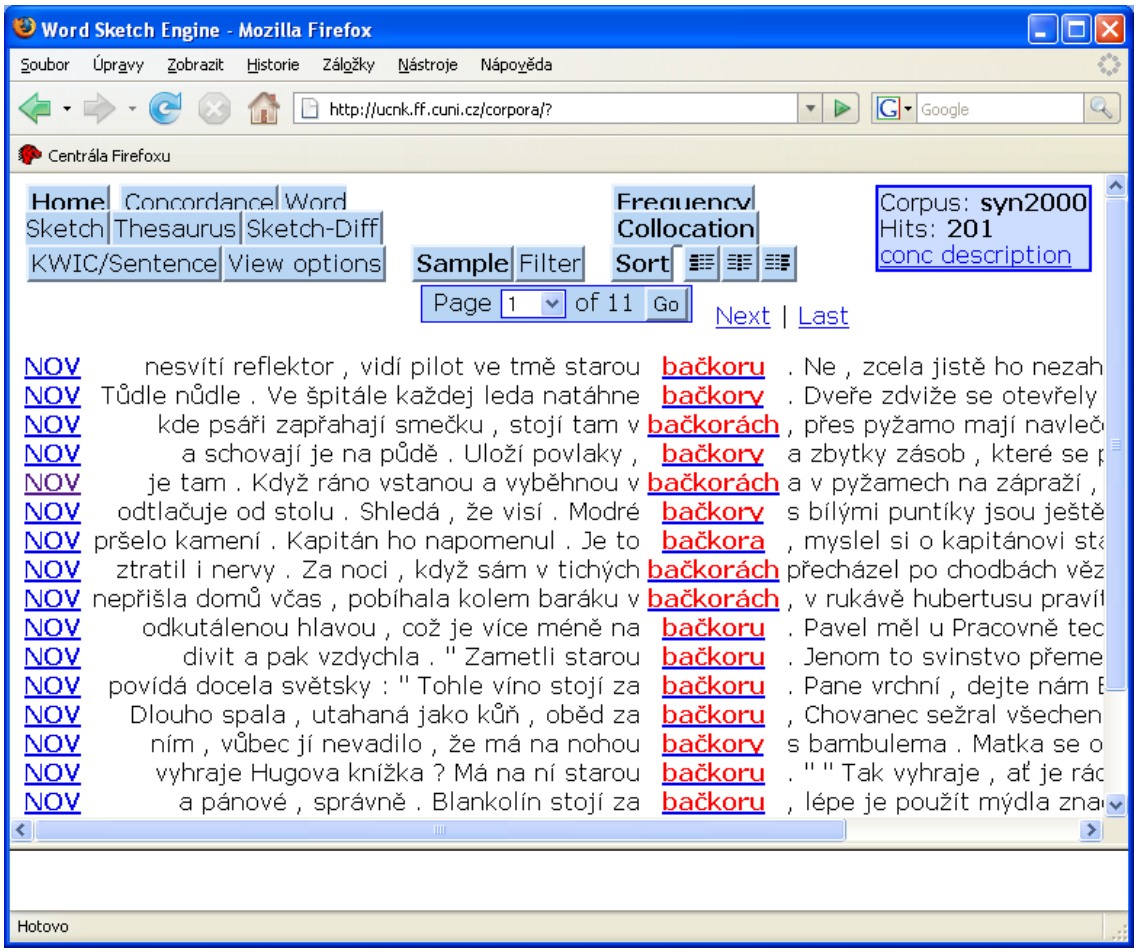
Do sekce „**Context**“ budeme zadávat informace pouze tehdy, jestliže chceme zpřesnit charakteristiku kontextu, ve kterém se námi hledané slovo má nacházet. Jestliže chceme kontext filtrovat.

Za vodícím ukazatelem „**Query Type**“ můžeme vybrat „**All**“, což bude fungovat jako pozitivní filtr (v Bonitu 1 P-filtr), který nám zobrazí pouze ty kontexty klíčového slova, které budou další (NÍŽE) zadané slovo (do vodícího ukazatele „Lemma“, jak vidno, zde nelze zvolit úroveň „Word“) obsahovat, a to v zadaném rozmezí (defaultně –5 vlevo (Left context) až 5 vpravo (Right context)). Vybereme-li „**None**“, kontexty klíčového slova, které další (NÍŽE) zadané slovo obsahují, se nám nezobrazí (v Bonitu 1 N-filtr).

Za vodícím ukazatelem „**Window Size**“ zadáváme velikost kontextu, ve kterém chceme, aby se další zadané slovo nacházelo. Ale pouze vždy buď vlevo od KWIC, nebo vpravo od něj. Funkce, kterou známe z Bonita 1 – tedy funkce, která mohla pracovat před i za KWIC současně, není v tomto formuláři k dispozici (srovnejte ale 4.4. Filtr !!). Zadáme-li totiž stejné slovo do obou polí (tedy „Left context“ i „Right context“), hledáme kontext, ve kterém se slovo vyskytlo JAK před, tak současně ZA hledaným klíčovým slovem (a to bývá málokdy). Dotaz na okolí musíme tedy vždy rozdělit na dvě části, levou a pravou (srovnejte ale 4.4. Filtr !!). Na druhou stranu nám tento způsob zpracování umožní relativně jednoduchým způsobem hledat kontext, kde před klíčovým slovem stojí jiné konkrétní slovo, než za ním. V Bonitu 1 jsme to dělali buď postupným použitím dvou P-filtrů za sebou, nebo v Dotaz-Grafické vytváření-Posloupnost, kde jsme mezi hledaná slova vkládali Opakovaně-Libovolná pozice.

1.4. Práce nad vyhledanou konkordancí

Poté, co jsme odeslali formulář do korpusu a co se nám zobrazila hledaná konkordance,



můžeme pokračovat v práci.

V pravém horním rohu vidíme informaci o tom, v jakém se právě nacházíme korpusu („**Corpus**“) a kolik bylo na zadaný dotaz odpovědí, tedy kolik je nalezeno výskytů („**Hits**“). V levé horní části nás tlačítka první řádky odvedou do jiných částí programu, začneme tedy řádkou druhou.

1.4.1. KWIC versus věta

V korpusové lingvistice je vyhledaná konkordance často zobrazována tak, že klíčové slovo (KWIC) je zarovnáno uprostřed, zvýrazněno barvou a kontext ubíhá vlevo a vpravo tzv. do nekonečna. Tlačítko „**KWIC/Sentence**“ nám umožňuje přepínat z tohoto způsobu zobrazení do zobrazení po celých větách (ty jsou pak zarovnány vlevo).

Možnost zobrazit si konkrétní kontext šíře nám ale zůstala, stačí, když levou myší jednou klikneme na KWIC, jehož širší kontext nás zajímá. Objeví se v dolním okně, podobně jako v Bonitu 1.

1.4.2. Možnosti zobrazení

Tlačítko „**View options**“ nás vede k formuláři, na kterém si můžeme vybrat, které hodnoty jednotlivých slov uvnitř textu, tedy „**Attributes**“ chceme zobrazit:

word	= samo slovo / konkrétní forma
lemma	= lemma slova
tag	= celý tag
lc	= slovo / konkrétní forma – ovšem nehledě na velikost mísmen

pos = tag slovního druhu

a některé značky duplicitně ještě jednou či samostatně (tyto jsou zavedeny kvůli zpracování Word Sketch-ů, toto zpracování totiž s korpusovými tagy pracovat nedokáže)

k = značka slovního druhu

g = jmenný rod

c = pád

Můžeme si tu ovšem vybrat pro zobrazení také hodnoty „**References**“, které chápeme jako vnější textu, stojící mimo něj a vztahující se k němu jako k celku:

token number = číselné vyjádření pozice slova v korpusu

doc. type = typ textu

doc.temp = rok vydání

doc.opus = značka jednoznačně identifikující text (a zařazená v seznamu textů)

Relativně málo užívané je zobrazení „**Structures**“, tedy strukturních značek textu (začátky vět, kapitol apod.):

doc = dokument (někdy kapitola románu, někdy celý román, v novinách většinou článek)

s = věta

Tyto možnosti se buď zobrazují jen pro klíčové slovo („**Display attributes**“ „**KWIC tokens only**“), nebo pro všechna slova („**For each token**“).

1.4.3. Vzorek

Tlačítko „**Sample**“ nás vede k formuláři pro „**Random sample**“, tedy náhodný výběr. Volíme zde množství řádek tohoto náhodného vzorku.

1.4.4. Filtr

S filtrem aplikovaným přímo při prvním zadávání dotazu jsme se setkali už v „3.2. Kontext“ a můžeme ho vidět na 3 obrázku v sekci „Context“.

Tlačítko „**Filter**“ nás vede k (v pořadí vlastně druhému) filtru, tento je aplikovaný po vyhledání konkordance; je totožný s funkcemi, které známe z Bonita 1: P-filtr ,zde „Filter positive“, a N-filtr, zde „Filter negative“. Pro velikost kontextu, ve kterém chceme filtr použít, platí stejná pravidla jako v Bonitu 1, levý kontext se zapisuje v záporných číslech (tedy např. –3, 0 = filtr bude aplikován do tří pozic vlevo od klíčového slova).

1.4.5. Jednoduché třídění – víceúrovňové třídění

Tlačítko „**Sort**“ nás vede k formuláři pro jednoduché třídění – „**Simple Sort**“ – které známe už z Bonita 1. Nenabízí sice možnost setřídit také klíčové slovo, což Bonito 1 umožňuje, zbylé funkce jsou však shodné. (Setřídění klíčového slova zleva je možné pomocí prostřední třídící ikony ve 4.6. – ovšem možnost třídění klíčového slova zprava /retrográdně/ Bonito 2 nenabízí vůbec).

„**Multilevel sort**“ je funkce, kterou Bonito 1 nemělo. Umožňuje totiž třídit „AŽ“ podle druhého, resp. třetího slova a slova bližší nebrat v úvahu. Podobně umožňuje také volit v rámci těchto pozic POSTUP třídění.

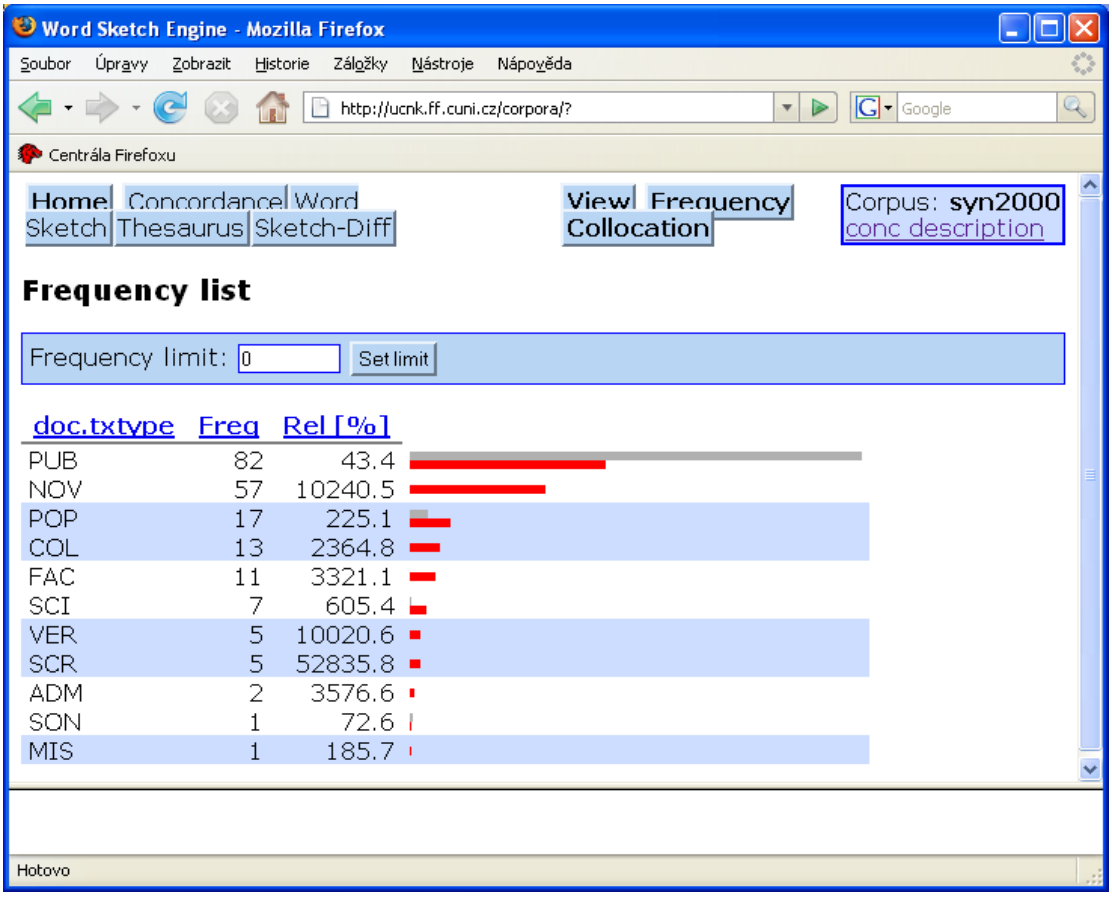
1.4.6. Jednoduché třídění ikonami

Následující tři tlačítka jsou ikonami pro rychle setřídění – vždy zleva: levého kontextu, klíčového slova, pravého kontextu.

1.5. Frekvence

Tlačítko „**Frequency**“ nás dovede k formuláři, který známe z Bonita 1 jako formulář pro „Frekvenční distribuci“, zde má název „**Multilevel frequency distribution**“. Na rozdíl od Bonita 1 nemůžeme jít dále za pozici 3 vlevo a v pravo od klíčového slova. Nemůžeme také sledovat frekvenci všech atributů, které jsme mohli sledovat v Bonitu 1 – můžeme sledovat: wordu / lc, lemmatu, tagu / pos, a dále: txtype, temp, opus.
Zajímá-li nás víceslovná jednotka, v Bonitu 2 (na rozdíl od Bonita 1) nemůžeme vytvářet statistiky uvnitř tohoto řetězce, protože se zaměřuje jen na okolní kontext. Z klíčového řetězce dokáže statisticky zpracovat pouze jeho první pozici.

Ve formuláři „**Text Type frequency distribution**“ si můžeme vybrat jeden z možných pohledů na distribuci klíčového slova: podle typů textů, podle roku vydání, podle jednotlivých děl. Poměr mezi absolutní frekvencí (zde „Freq“), a relativní frekvencí (tedy frekvencí přepočtenou s ohledem na poměrnou velikost kategorie zde „Rel“) v konkrétním námi zvoleném pohledu je tu znázorněn i sloupcovým grafem.



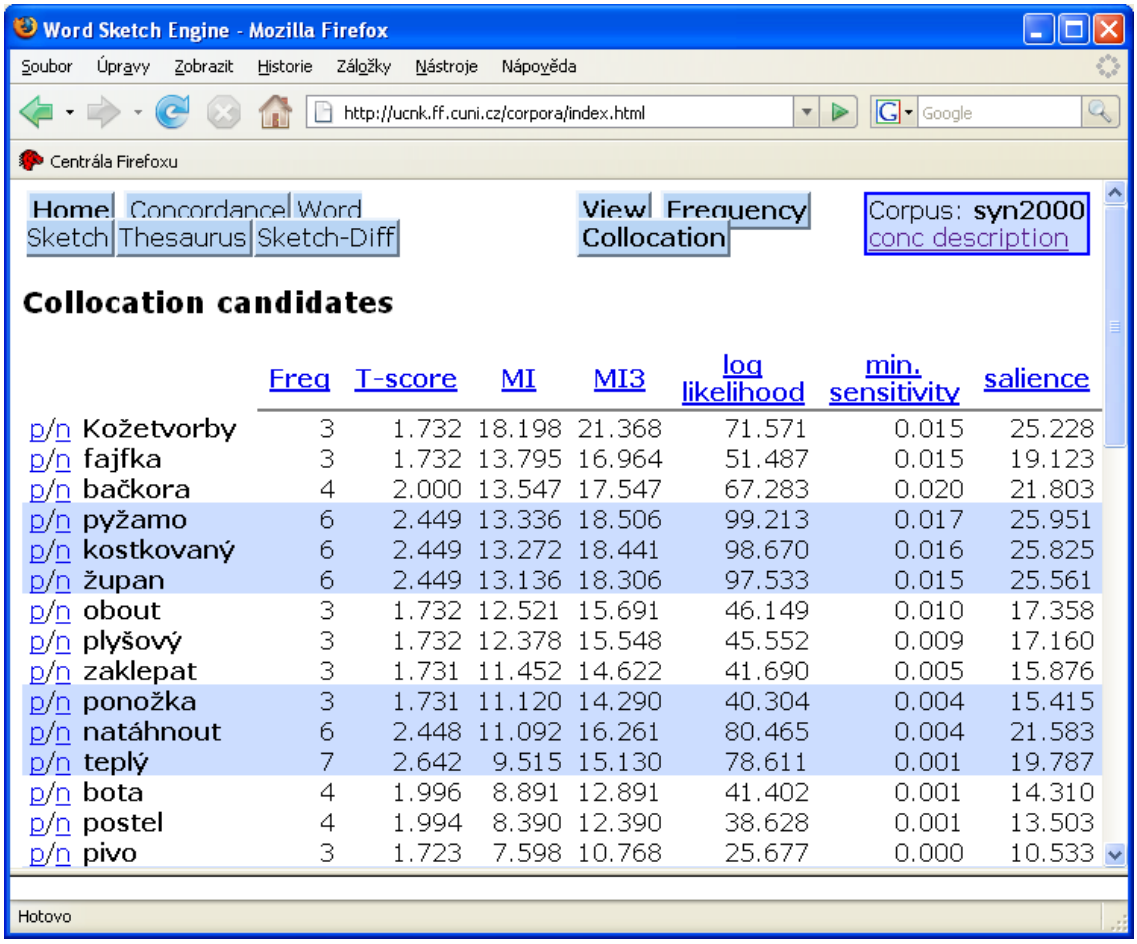
Sloupcový graf je konstruován následovně:
položky jsou řazeny za sebou podle klesající absolutní frekvence („Freq“). V celé tabulce sloupcového grafu se v souvislosti s tím proporcionálně mění jen zobrazení ČERVENÉ absolutní frekvence.
ŠEDÁ relativní frekvence (přepočtená na základě poměrné velikosti kategorie) je znázorněna proporcionálně pouze vzhledem k sobě odpovídající červené hodnotě absolutní frekvence.
Zajímavou novinkou oproti Bonitu 1 je tu možnost nechat si zobrazený seznam „přetřídit“ (abecedně podle kolokujících slov / lemmat, nebo podle relativní frekvence), k tomu slouží aktivní horní popisky sloupců, na které stačí kliknout.

1.6. Kolokace

Tlačítko „**Collocation**“ nás vede k formuláři „Collocation candidates“, který odpovídá v Bonitu 1 funkci Statistiky-Kolokace. I zde si vybíráme, na čem chceme statistiky počítat (_Attribute“: word – lemma – tag – lc – pos – k – g – c), v jak velkém kontextu se bude kolokace zjišťovat („In the range from“). Jaká má být minimální frekvence započítávané jednotky v korpusu („Minimum frequency in corpus“) a jaká má být minimální frekvence započítávané jednotky v námi stanovené velikosti kontextu („Minimum frequency in given range“). Vodící ukazatel „Maximum number of displayed lines“ poukazuje na možnost zvolit si velikost seznamu, který se z korpusu spočítá.

Podobně jako v Bonitu 1 mají zobrazené sloupce scorů aktnivní horní popisky sloupců, takže chceme-li získaný seznam přetřídit podle jiného scoru, stačí na daný horní popisek kliknout. Na rozdíl od Bonita 1 jedna je tu ale 6 statistických měr, podle kterých můžeme nechat kolokace setřídit („**Sort by**“), a rozdíl je i v tom, že si můžeme vybrat, které z možných měr se nám ve výsledné tabulce zobrazí („Show functions“).

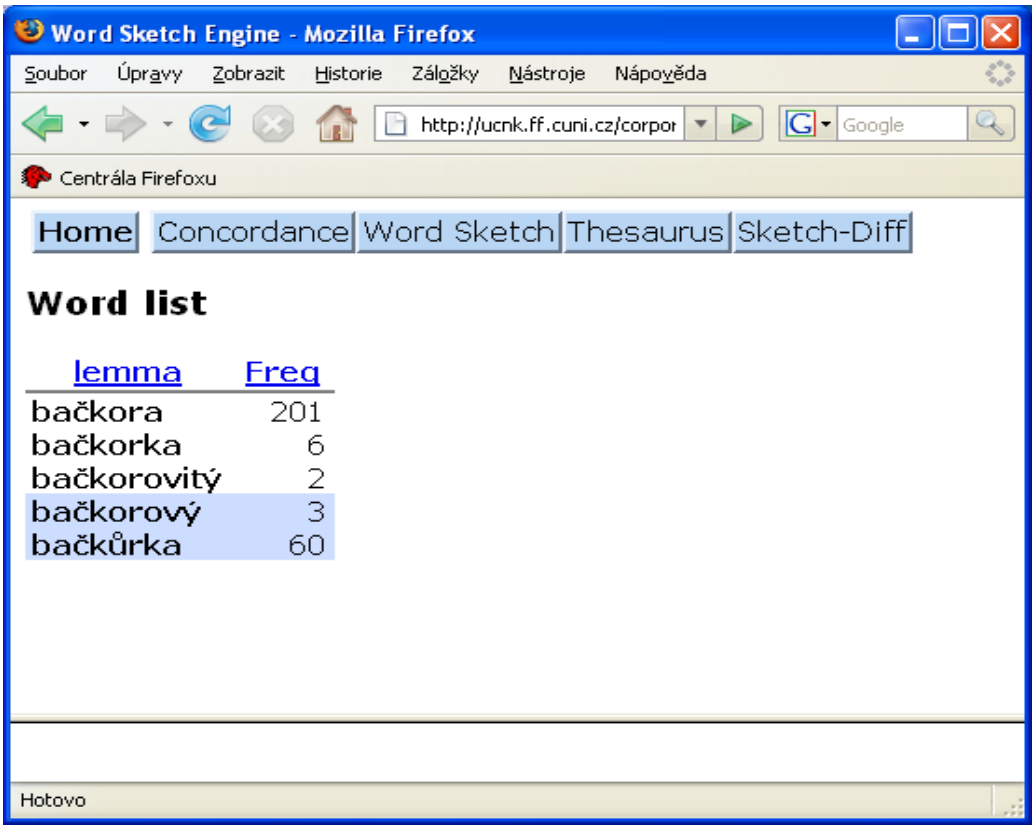
Na rozdíl od Bonita 1 tu máme velmi zajímavou možnost jednoduše se na kolokáty ze zobrazeného seznamu podívat přímo do korpusu. Zcela vlevo máme funkce **p**/**n** , které jsou shodné s P-filtrem a N-filtrem Bonita 1. Klikneme-li tedy na P, zobrazí se kolokace daného slova (ze seznamu) s naším klíčovým slovem a to podle hodnot, které jsme si zadali do formuláře Kolokace (tedy např. pokud je v rozmezí, které jsme zadali, pokud je v korpusu a v okolí KWICu v počtech, které jsme zadali).



Corpus

2.1. Seznam slov

Pod vodícím ukazatelem "**Word list**" se skrývá funkce, kterou známe z Bonita 1 jako Seznam slov. V Bonitu dvě je to funce vyčleněná ze sady funkcí "Concordance" ven – musíme tedy znovu zadat korpus, ve kterém chceme pracovat. Podobně jako v Bonitu 1 můžeme zadávat úroveň, kterou prohledáváme ("Attribute": word, lc, lemma, tag, pos, k, g, c), minimální frekvenci hledané jednotky ("Minimum frequency in corpus") a maximální počet zobrazených jednotek ("Maximum number of displayed lines"). Zobrazený výsledek můžeme – podobně jako v Bonitu 1 – přetřídit podle frekvence (klinutím na "Freq"), nebo zpět podle abecedy (kliknutím na "word").
Na rozdíl od Bonita 1 ale nefunguje druhé kliknutí. V Bonitu 1 totiž toto druhé kliknutí setřídilo kategorii "zpětně" – tedy od nejnižší frequence k nevyšší, od konce abecedy k začátku (ovšem i tentokrát prográdně).
Na rozdíl od Bonita 1 také nelze v tomto zobrazení vybrat jednu položku a nechat si zobrazit její korpusové konkordance.



2.2. Vytvoření subkorpusu

Vodící ukazatel "**Create subcorpus**" známe z Bonita 1 jako Vytvoření subkorpusu. V Bonitu 2 klikneme na ukazatel a ve formuláři jako první krok zadáme korpus, ze kterého chceme subkorpus vytvářet. Pak už jen vyplníme jméno subkorpusu ("New subcorpus name") a zaškrtneme v kategoriích, ve kterých chceme vytvářet

typy textu	= doc.txtype
rok	= doc.temp
konkrétní dílo	= doc.opus

konkrétní hodnoty. V pravém sloupci vidíme velikost dané kategorie, vyjádřenou množstvím slov.

2.3. Zrušení subkorpusu

Jednoduchý a názorný formulář se nám objeví pod "**Delete Subcorpora**".

Sketches

3.1. Nárys užívání

I v češtině často používáme nepřeložené označení "**Word sketch**", které zastupuje relativně komplexní funci, která spojuje následující úkony:

1) spočítání absolutní frekvence kolokace
2) spočítání hodnoty salience (MI-score krát logaritmus frekvence = tedy nové MI-score, které je upraveno tak, aby se velmi nízké frekvence poněkud znormalizovali) a setřídění podle této hodnoty
3) rozdělení seznamu podle syntaktických vztahů (uvnitř kterých zůstává setřídění podle hodnoty salience) /seznam následuje/.

Podobně jako u Seznamu slov, musíme i zde na začátku zadat korpus, ve kterém chceme pracovat. Dále zadáme "Lemma", které nás bude zajímat, možnost pracovat na úrovni "Word" tu nemáme.
"Sort grammatical relations" bylo snad původně zamýšleno jako funkce, která – pokud ji NEzvolíme – spočítá hodnoty kolokací, ale nebude je dělit podle syntaktických vztahů. Funkce není aktivní a pokud chceme mít nedělený seznam setříděný podle salience, použijeme v "Concordance" po vyhledání klíčového slova funkci "Collocation".
"Minimum frequency" je zde myšleno jako minimální frekvence v okolí KWIC-u, se kterou se bude ještě pracovat, nikoli tedy v minimální frekvence v celém korpusu (tu zadávat nemůžeme, na rozdíl od Bonita 1).
"Minimum salience" je minimální hodnota salience, která se bude ve výsledcích zobrazovat.
"Maximum number of items in a grammatical relation" je maximální počet jednotek zobrazených v jedné kategorii (syntaktických vztahů).

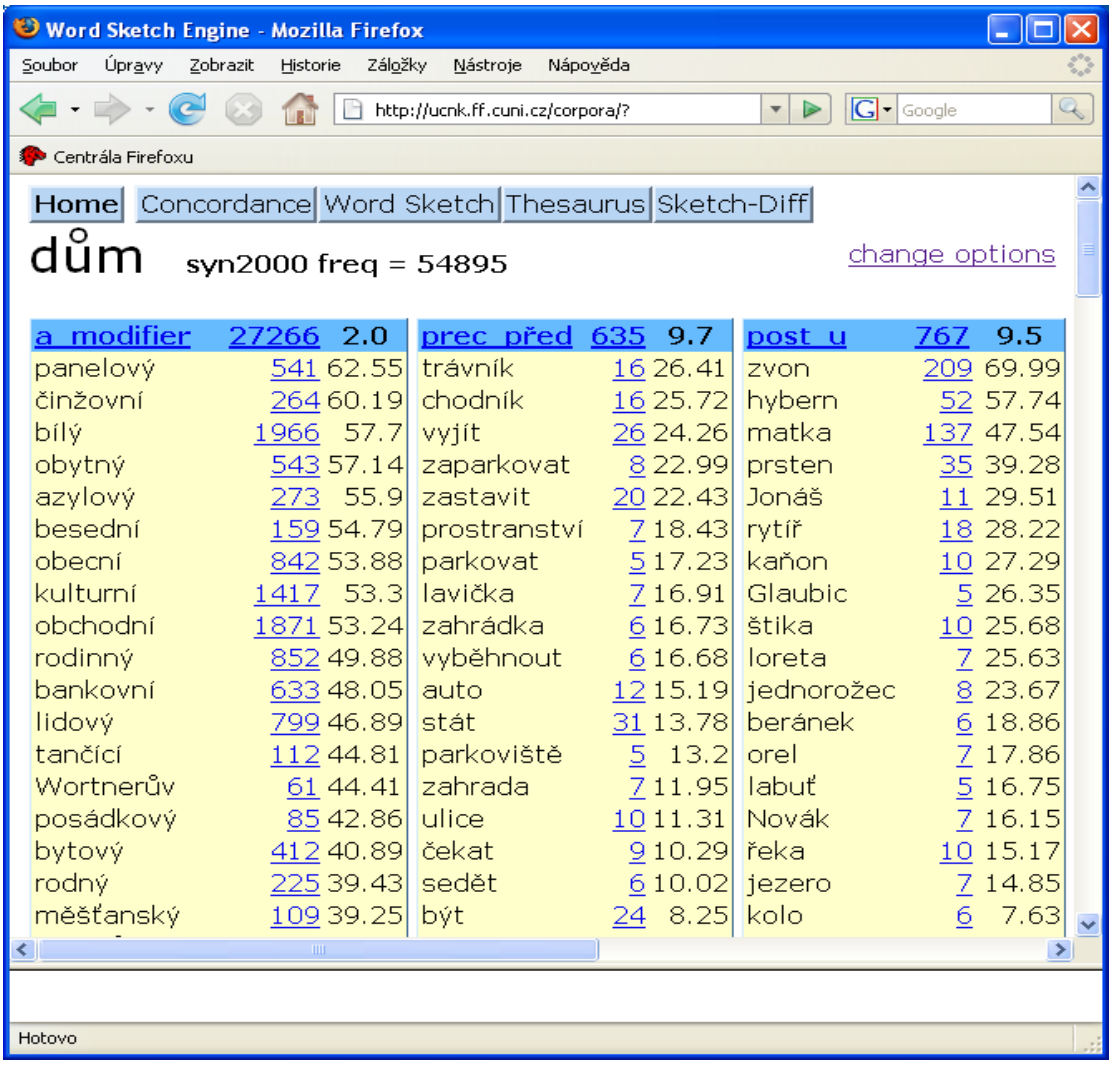
Rozmezí (velikost okna), ve kterém jsou kolokace počítány, je –5 / 5 a nelze bohužel změnit.

Nejčastěji zobrazované syntaktické kategorie jsou:

a modifier	přívlastek (shodný) stojící vlevo (od KWIC)
	panelový dům
prec X	X = konkrétní předložka, např. před
prec "před"	předložkový vztah se substantivem vlevo (od KWIC)
	stejně jmenný jako slovesný
	trávník před domem
	vyběhnout před dům
post X	X = konkrétní předložka, např. s
post "s"	předložkový vztah se substantivem vpravo (od KWIC)
	stejně jmenný jako slovesný
	dům se zahradnou

gen 1	genitivní vztah 2 substantiv (KWIC je v nominativu) Dům kultury ...
gen 2	genitivní vztah 2 substantiv (KWIC je v genitivu) majitel domu
prec verb	verbum vlevo stavějí domy
post verb	verbum vpravo dům chátrá
post inf	infinitiv vpravo dům prodat se běží podívat
byt adj	vztah s pomocným slovesem "být" a predikátovým adjektivem (prvkem tagovaným jako adjektivum) dům je prádný
prec prep	předložka vlevo před domem
post prep	předložka vpravo běžel ke (dvěřím)
has subj	"má podmět (jaký)" vztah verbálního KWIC-u a substantiva v nominativu lhůta běží
is subj of	"je podmětem pro" vztah KWIC-u v nominativu a verba dům vyhořel
is obj2 of	"je předmětem ve 2 pádě pro" vztah KWIC-u v genitivu a verba se netýká domu
is obj3 of	"je předmětem ve 3 pádě pro" vztah KWIC-u v dativu a verba šel domu **!! chybně tagováno**
has obj4	"má (tento) předmět ve 4 pádě" vztah verbálního KWIC-u a substantiva v akuzativu běžel maraton
is obj4 of	"je předmětem ve 4 pádě pro" vztah KWIC-u v akuzativu a verba prohledali dům
has obj7	"má (tento) předmět v 7 pádě" vztah verbálního KWIC-u a substantiva v instrumentálu běží ulicí
is obj7 of	"je předmětem v 7 pádě pro" vztah KWIC-u v instrumentálu a verba prochází domem
coord	vztah lexikálně vyjádřené koordinace s jednotkou stejné třídy dům a zahradu se otočil a běžel

Výhodnou a jistě často užívanou funkcí bude možnost okamžitě se podívat na konkrétní kolokát, který nás zaujal. stačí, když klikneme na aktivní absolutní frekvenci kolokace



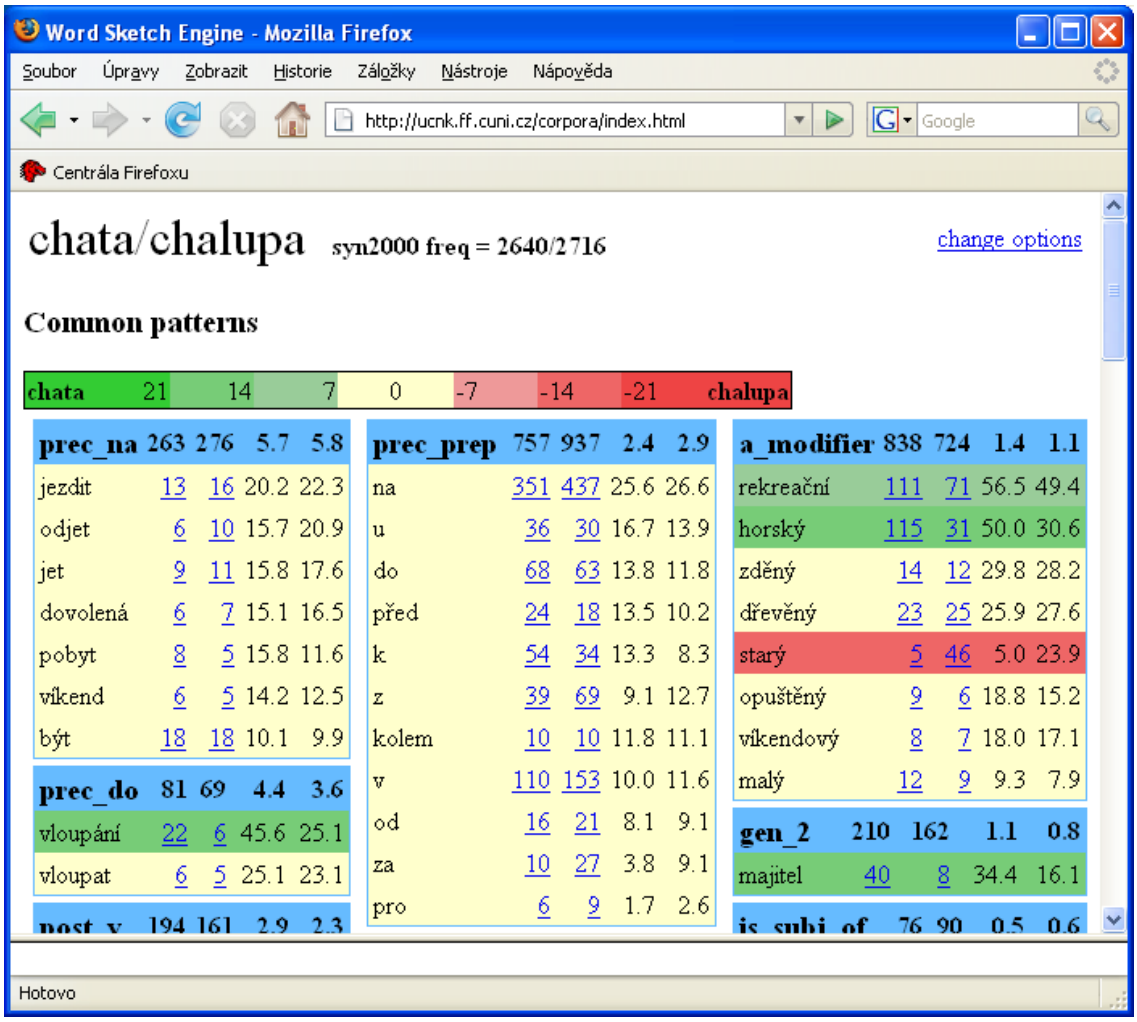
a_modifier	27266	2.0	prec_před	635	9.7	post_u	767	9.5
panelový	541	62.55	trávník	16	26.41	zvon	209	69.99
činžovní	264	60.19	chodník	16	25.72	hybern	52	57.74
bílý	1966	57.7	vyjít	26	24.26	matka	137	47.54
obytný	543	57.14	zaparkovat	8	22.99	prsten	35	39.28
azylový	273	55.9	zastavit	20	22.43	Jonáš	11	29.51
besední	159	54.79	prostranství	7	18.43	rytíř	18	28.22
obecní	842	53.88	parkovat	5	17.23	kaňon	10	27.29
kulturní	1417	53.3	lavička	7	16.91	Glaubic	5	26.35
obchodní	1871	53.24	zahrádka	6	16.73	štika	10	25.68
rodinný	852	49.88	vyběhnout	6	16.68	loreta	7	25.63
bankovní	633	48.05	auto	12	15.19	jednorožec	8	23.67
lidový	799	46.89	stát	31	13.78	beránek	6	18.86
tančící	112	44.81	parkoviště	5	13.2	orel	7	17.86
Wortnerův	61	44.41	zahrada	7	11.95	labuť	5	16.75
posádkový	85	42.86	ulice	10	11.31	Novák	7	16.15
bytový	412	40.89	čekat	9	10.29	řeka	10	15.17
rodný	225	39.43	sedět	6	10.02	jezero	7	14.85
měšťanský	109	39.25	být	24	8.25	kolo	6	7.63

3.2. Rozdíly v užívání 2 jednotek

Tato funkce, tedy "**Word sketch difference**" zpracovává způsobem popsaným v 3.1. dvě jednotky, jejich výsledky porovnává a zobrazuje barevně odlišeně (první položka je zelená a různý stupeň přináležitosti k ní je naznačen tónem/sytostí barvy, druhá položka je červená, stupeň přináležitosti je řešen stejně). Tento způsob je zřetelně vidět na tabulce v záhlaví výsledků, kde je vyznačeno i střední, nespecifické pásmo krémové barvy a hraniční hodnoty odlišných tónů/sytostí barev.

chata	21	14	7	0	-7	-14	-21	**chalupa**

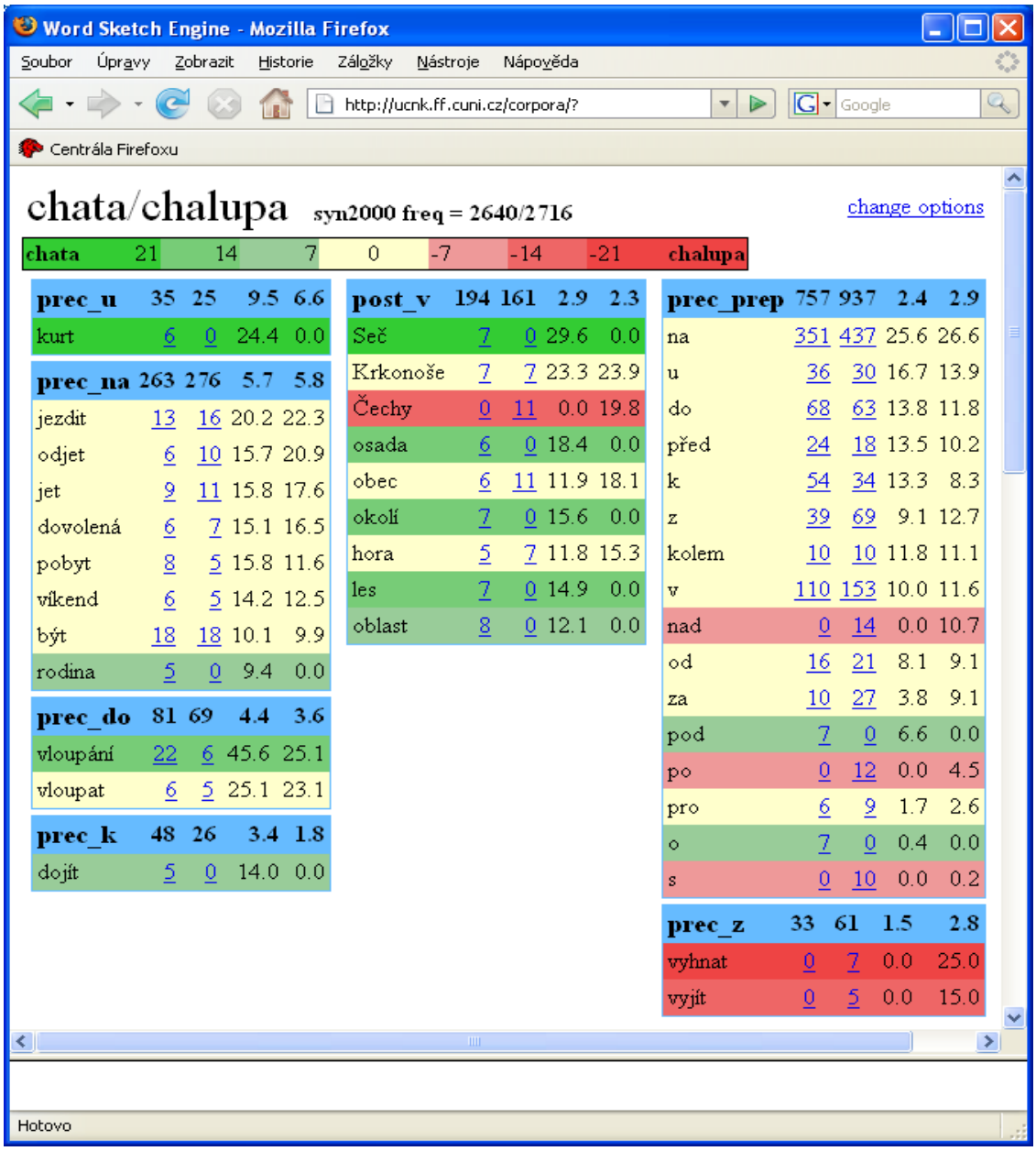
Celkové zobrazení je ovšem celkem názorné. Má dvě varianty. V Bonitu 2 je předem zaškrtnuta za vodícím ukazatelem "**Separate blocks**" položka "common/exclusive blocks". Toto funkce nám vytvoří zobrazení:
– jedno pole pro společné jednotky
– další dvě pole pro jednotky specifické pro každé z klíčových slov



Protože vedle pole pro jednotky společné jsou zde další dvě pole, máme dvojí možnost zadat množství zobrazovaných jednotek: jednak v poli společném "Maximum number...of the common block", jednak v polích disjuktních "Maximum number...of the exlusive block".

Pokud ovšem za vodícím ukazatelem "**Separate blocks**" zvolíme položku "all in one block", pak budete mít všechny jednotky statistikou považované za relevantní v jediném společném přehledu.

Pozor!!!
Barva tu nesouvisí s výlučností pro jedno z klíčových slov, souvisí pouze s hodnotou salience, která může být i pro slova (v dané dvojici) specifická pro jedno z klíčových slov relativně nízká.
Srovnejte v tabulce níže např. kolokát „oblast“.



3.3. Tezaurus

Za vodícím ukazatelem „**Thesaurus**“ najdeme formulář pro hledání lemmat (v konkrétním korpusu, který si musíme nejdříve vybrat), v jejichž okolí se vyskytuje určitý počet stejných jednotek jako v okolí zadaného klíčového slova. Jsou tedy porovnány výsledky „Sketchů“ pro celý korpus a výsledný seznam obsahuje ty jednotky, u kterých byla shledána podobnost s klíčovým slovem („Minimum similarity between cluster items“) vyšší, než ve formuláři zadaná (nebo jí rovná).

Jak je tato „podobnost“ počítána, není jasné. Z výsledků není zřejmé bohužel ani to, k čemu referují hodnoty zobrazované za každým uváděným lemmatem. Lze jen říci, že se výsledky skutečně proměňují v závislosti na

- počtu zobrazovaných lemmat („Maximum number of items“)
- minimální podobnosti s klíčovým slovem („Minimum similarity between cluster items“)

Za zmínku zde stojí organizace (shlukování) zobrazovaných lemmat. Lemmata nejsou zobrazována podle jim odpovídajících hodnot (jejichž referenci bohužel neznáme), ale podle vzájemných podobností. V tezauru pro slovo „dům“ tedy můžeme vidět následující shluky – vybíráme:

budova	0.382	byt 0.361 objekt 0.295 prostor 0.201 stavba 0.2 zařízení 0.182 areál 0.162
město	0.261	obec 0.194 země 0.188 Praha 0.157 republika 0.155
místnost	0.191	pokoj 0.172 hala 0.166 sál 0.152
obchod	0.175	služba 0.165
rodina	0.175	dítě 0.154
majetek	0.161	pozemek 0.151