

PRÁCE S PRALEDEM A STRUKTURA HESLA V LDB PRALEX

Jindra Světlá – Martina Habrová, ÚJČ AV ČR, v. v. i., oddělení současné lexikologie a lexikografie

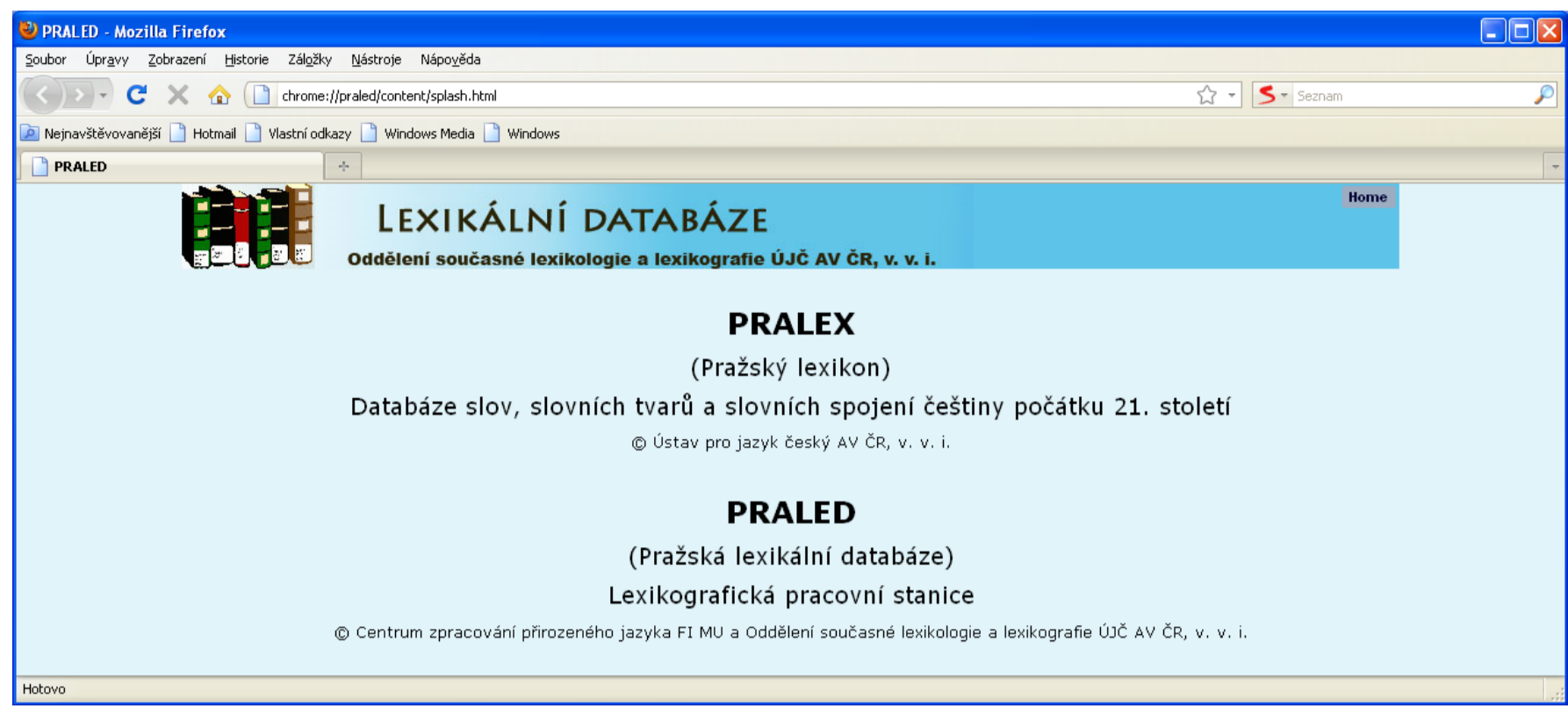
ÚVOD

Cíl výzkumného záměru ÚJČ AV ČR, v. v. i.: Vytvoření *datové lexikální zásoby českého jazyka počátku 21. století*
Základní úkol LDB Pralex: shromáždění a rozřídění jazykových materiálů dokládajících především dnešní úzus
Celkové uspořádání databáze: abecední (hesla s velkými a malými počátečními písmeny lze vyhledat společně nebo samostatně)

1 NA POČÁTKU BYL PRALED

Co představuje název PRALED:

- software pro databáze zpracování
- prázdné jádro databáze s nástroji pro lexikografický popis české slovní zásoby
- lexikografická pracovní stanice
- zkratka z označení PRAžská LEXikální Databáze



obr. 1: Vstupní obrazovka LDB Pralex

Funkce PRALEDU:

- zakládání nových hesel, jejich mazání nebo duplikování
- vyhledávání (jednoduché – pokročilé vyhledávání)
- tisk různých sestav
- přístup k DEBDictu
- hypertextové propojení jednoslovných a vícenoslovných hesel

Práce se heslovým formulářem:

- zápis a ukládání dat
- opravy, doplňování a mazání dat
- duplikování dat
- zobrazování dat v příslušném nástroji a v náhledu hesla
- prohlášení vložených údajů – různé náhledy hesla

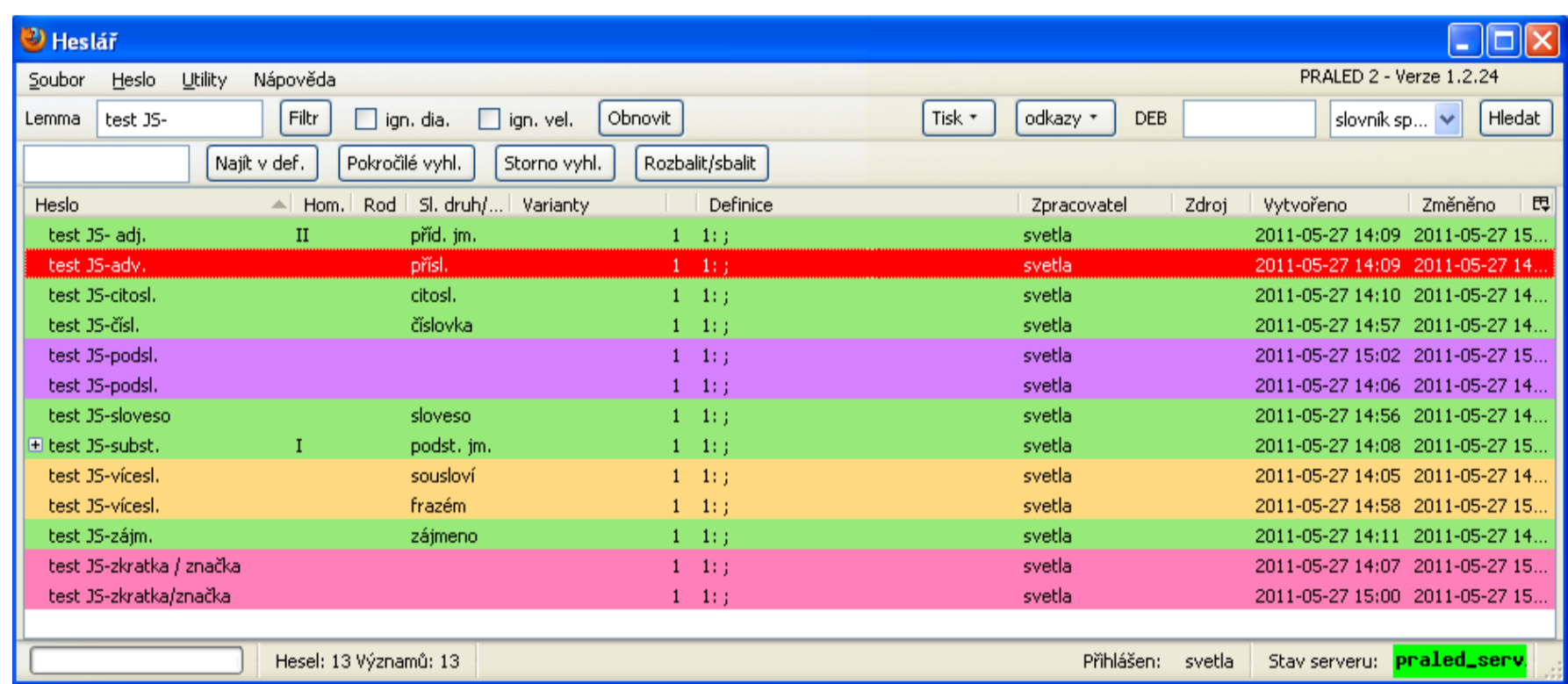
Práce se SEZNAMY:

- Některá hesla propojujeme mezi sebou pomocí tzv. Seznamů, které umožňují již v průběhu prací realizovat hypertextové odkazy (prokliky), a tudíž i snadné přechody mezi hesly, otevírající více hesel najednou na ploše obrazovky a jejich paralelní zpracování, vyhodnocování a současně revize či prohlášení.
- Pomocí **hypertextových odkazů** umístěných v jednotlivých seznamech – lze zachytit nejrůznější systémové vztahy, např. homonymii, variantnost, synonymii či antonymii, i vztahy mezi jednoslovnými a vícenoslovnými hesly (frázemi či souslovími a jejich komponenty), stejně tak jako vztahy mezi apelativy a proprii, nebo mezi zkratkami a jejich plnými tvary zpracovými jako samostatná hesla.

2 STRUKTURA LEXIKÁLNÍ DATABÁZE

1) MAKROSTRUKTURA (struktura hesláře)

- hesla jednoslovná
- hesla vícenoslovná
- zkratky a značky
- podslovní části



obr. 2: Typy hesel v Hesláři

2) MIKROSTRUKTURA (struktura hesla)

Součástí heslového formuláře:

- Karta HESLO
- Karta VÝZNAM
- Přidávání / odebrání významů (ouška karet)
- Okna s nástroji a položkami (tačítka, nabídky v menu, pole k vyplnění nebo zaškrtnutí)
- Seznamy pro hypertextové odkazy a propojování hesel

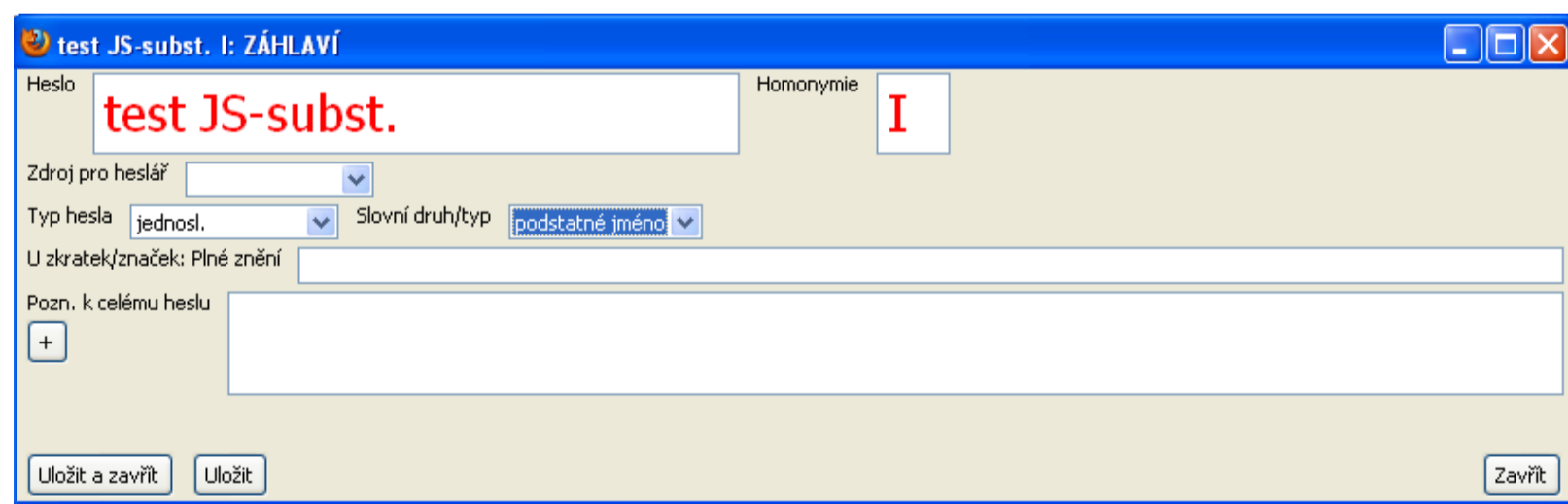
Nástroje a položky formuláře pro zápis dat:

- Karta HESLO obsahuje tyto nástroje: ZÁHLAVÍ – ADMIN. – VÝSL. – PŮVOD – DÉLENÍ – ETYMOL. – STAT. – GRAM. – VAR. – SLOVOTVORBA; EXPR. – ÚZEMNÍ PŘÍZNÁK – DOBOVÝ PŘÍZNÁK – STYLOVÝ PŘÍZNÁK – OBOR – OBLAST UŽITÍ – KPI – SUBSTANDARD (příznakové zatím až na výjimky neurčujeme, toto je orientačně připraveno pro slovník);
- SEZNAMY platné pro celé heslo (Seznam zkratk – Seznam sousloví – Seznam frázem – Seznam komponentů u vícenoslovných LJ – Obecná/vlastní jména – Seznam homonym) – slouží pro hypertextové odkazy a propojování celých hesel (lexemů)

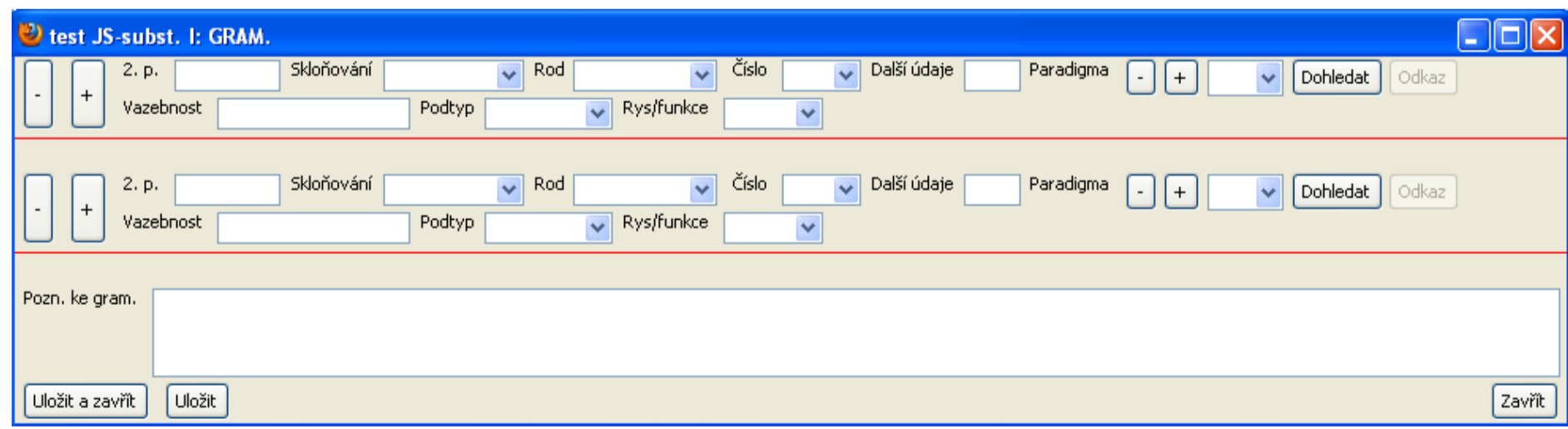
Povinné vyplňované položky na kartě Heslo (totožné pro všechny slovní druhy):

- Lemna (Heslo)
- Homonymie (v relevantních případech)
- Zdroj pro heslář
- Typ hesla
- Slovní druh
- Zpracovatel
- Statistika (v relevantních případech)
- Varianty (v relevantních případech)
- Seznamy (v relevantních případech)

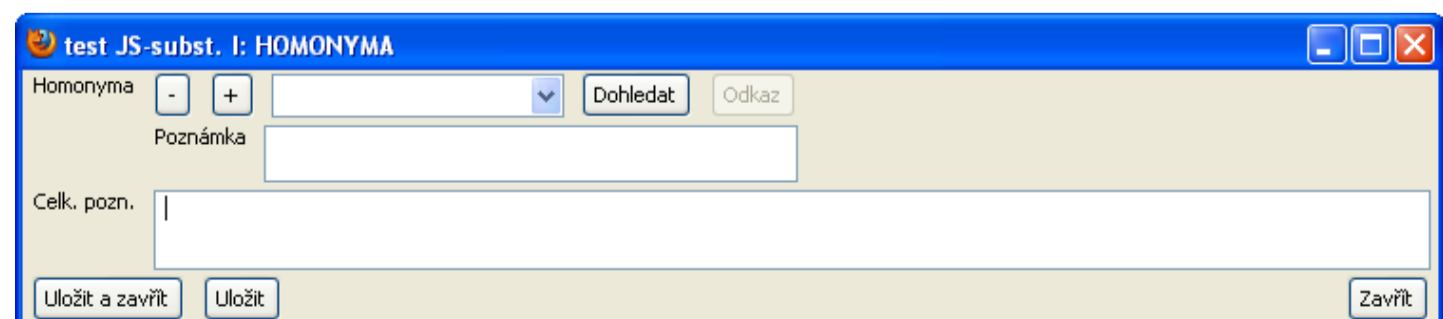
UKÁZKY NÁSTROJŮ POUŽÍVANÝCH V RÁMCI VÝZKUMNÉHO ZÁMĚRU na kartě Heslo:



obr. 3: Nástroj ZÁHLAVÍ na kartě Heslo



obr. 4: Nástroj GRAM. (na kartě Heslo i Význam)



obr. 5: Seznam homonym

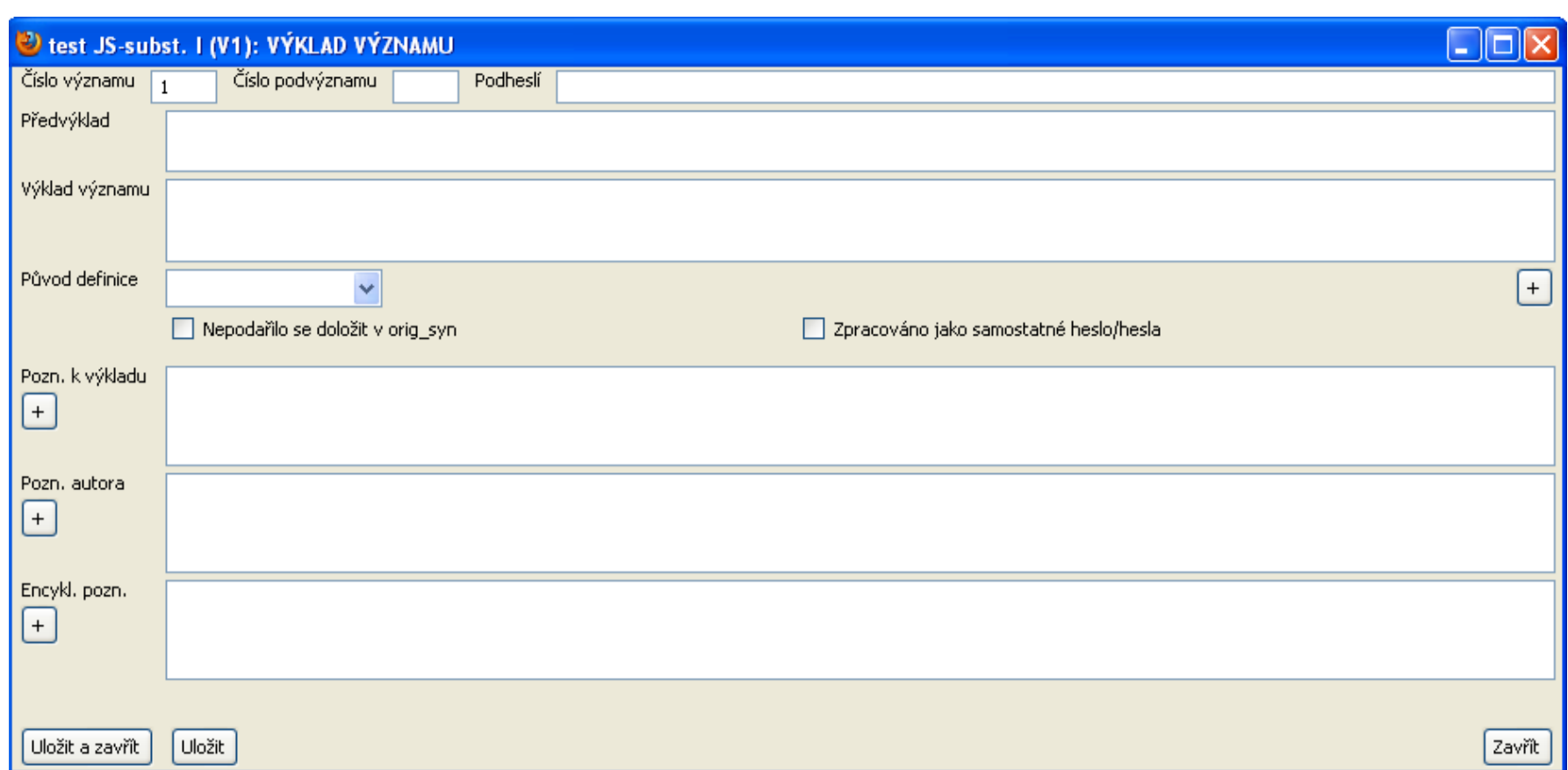
- Karta VÝZNAM obsahuje tyto nástroje: GRAM. – PŮVOD – EXPR. – ÚZEMNÍ PŘÍZNÁK – DOBOVÝ PŘÍZNÁK – STYLOVÝ PŘÍZNÁK – OBOR – OBLAST UŽITÍ – KPI – SUBSTANDARD (příznakové zatím až na výjimky neurčujeme, toto je orientačně připraveno pro slovník); VÝKLAD VÝZNAMU – EXEMPLIFIKACE – SLOVOTVORBA – NOVÉ DOKLADY
- Nástroj VÝKLAD VÝZNAMU (totožný pro všechny slovní druhy a pro všechny založené významy)

Povinné položky:

- Výklad významu
- Původ definice
- Pozn. k výkladu
- Podheslí (v relevantních případech)
- Nepodařilo se doložit (v relevantních případech)
- Zpracováno jako samostatné heslo (v relevantních případech)

+ fakultativně:

- Předvýklad
- Pozn. autora



obr. 6: Nástroj VÝKLAD VÝZNAMU

- SEZNAMY platné pro jednotlivé významy (Seznam synonymum, antonym, hyperonym, hyponym, kopolyonym, seznam sémů – většinu z nich zatím nepoužíváme) – slouží pro hypertextové odkazy a propojování na úrovni jednotlivých lexémů
- Nástroj EXEMPLIFIKACE

Povinné položky:

- upravené doklady (v relevantních případech)
- větné doklady – vždy

+ fakultativně:

- blok V názvech
- Pozn. k E celkové

- Nástroj **Nové doklady** používáme v relevantních případech jako speciální blok Exemplifikace.

3 PRÁCE V LDB PRALEX

Co představuje název PRALEX:

- obsah lexikální databáze (databáze slov, slovních tvarů a slovních spojení)
- základní zpracování jednotlivých položek (jádro budoucích slovníkových hesel)
- východisko pro nový komplexní popis slovní zásoby současné češtiny (LDB Pralex chápeme jako přípravnou fázi budoucího slovníku)
- východisko pro další teoretický výzkum slovní zásoby
- heslář jako strukturálně uspořádaný celek (inventář jazykových jednotek) a východisko pro tvorbu různých slovníků
- zachycení systémových vztahů ve slovní zásobě (hypertextové odkazy)
- materiálová základna pro nový výkladový slovník
- zkratka z označení PRAžský LEXikon

Zpracování heslových slov (po jednotlivých slovních druzích):

Zpracované položky LDB obsahují

- informace o registraci výskytu daného lemmatu a jeho zpracování v předchozích výkladových slovnících
- určení slovního druhu nebo typu hesla
- pracovní definici (zatím jen podle předchozích výkladových slovníků, hlavně s oporou o ssjč)
- korpusové doklady na konkrétní užití utištěné podle typů spojitelnosti syntaktické, lexikální a sématické
- autorské poznámky a komentáře

U každého zpracovávaného databázového hesla uvádíme

a) údaje platné pro celé heslo:

Lemna: Lemna odpovídá reprezentativnímu tvaru u slovníkových hesel, ale v databázi se v poli pro lemna uvádí zásadně jediný tvar.

Homonymie:

Homonymie značíme římskou číslicí a pracovně (po potřebě LDB) je číslujeme v pořadí podle slovních druhů, počítaje substantivem (všechna homonymia v rámci dané řady rovněž hypertextově propojujeme mezi sebou, aby se dala prohlížet zároveň). Všechny typy homonym (morfologická, slovnodruhová, slovnovárná i lexikální) pracovně uvádíme jako samostatné položky databáze. (Homonymia hypertextově propojujeme mezi sebou pomocí seznamu.)

Typ hesla:

viz makrostruktura (heslo jednoslovné – vícenoslovné – podsl. část – zkratka/značka)

Slovní druh/typ:

Zařazení hesla ke slovnímu druhu až na některé výjimky zatím zachováváme v souladu se SSJČ, případně jen komentujeme difference či novější pojetí podle Slovníku spisovné češtiny pro školu a veřejnost (SSJČ). K určení **typu lexikální jednotky** (dále LJ) patří rozhodnutí, zda jde v našem pojetí o sousloví, frázem, citátový výraz, vlastní jméno, zkratku nebo značku.

Variantnost:

V LDB Pralex pracovně registrujeme jako samostatné databázové položky téměř všechny variantní podoby uvedené v SSJČ a dalších zdrojích. Všechny informace uváděné v rámci jednoho hesla se musí vztahovat k dané variantní podobě lemmatu. Všechny varianty u každého hesla zapisujeme na seznam variant a hypertextově propojujeme mezi sebou.

Zdroj pro heslář:

Tento údaj je pro nás pracovní informací o tom, odkud jsme dané lemna (z hlediska formy) převzali do našeho hesláře. Uvedený údaj má dvě části, protože vycházíme z konfrontace korpusových a vybraných referenčních zdrojů.

Gramatické charakteristiky:

Sekci gramatických informací vyplňujeme zatím jen v případě potřeby, ale v rámci slovníkových pracovně budou pochopitelně doplněny všechny informace, které se zde tradičně uvádějí – příslušná pole jsou připravena včetně výběru z nabídek, které popis usnadňují.

b) údaje platné pro jednotlivé významy:

Výklad významu:

Zatím jde většinou jen o přiřazení **pracovní výkladové definice** na základě SSJČ, popř. dalších stávajících slovníků, k jednotlivým významům, v případě potřeby v jejím přechodnosti či komentování. K tomu máme k dispozici bohatý poznámkový aparát.

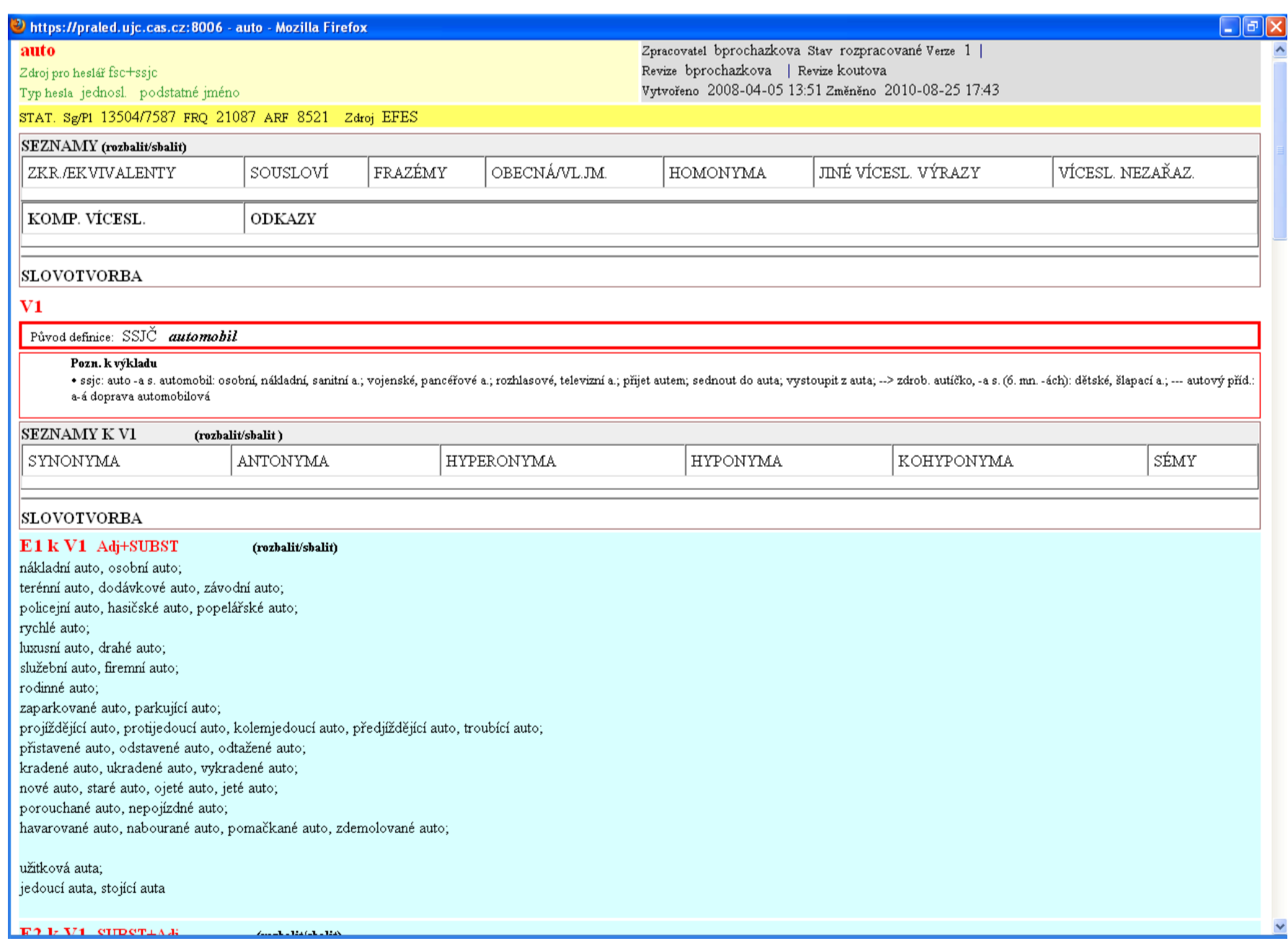
Poznámkový aparát (na různých úrovních hesla):

V jednotlivých poznámkách evidujeme nejrůznější informace, které později využijeme při slovníkovém zpracování.

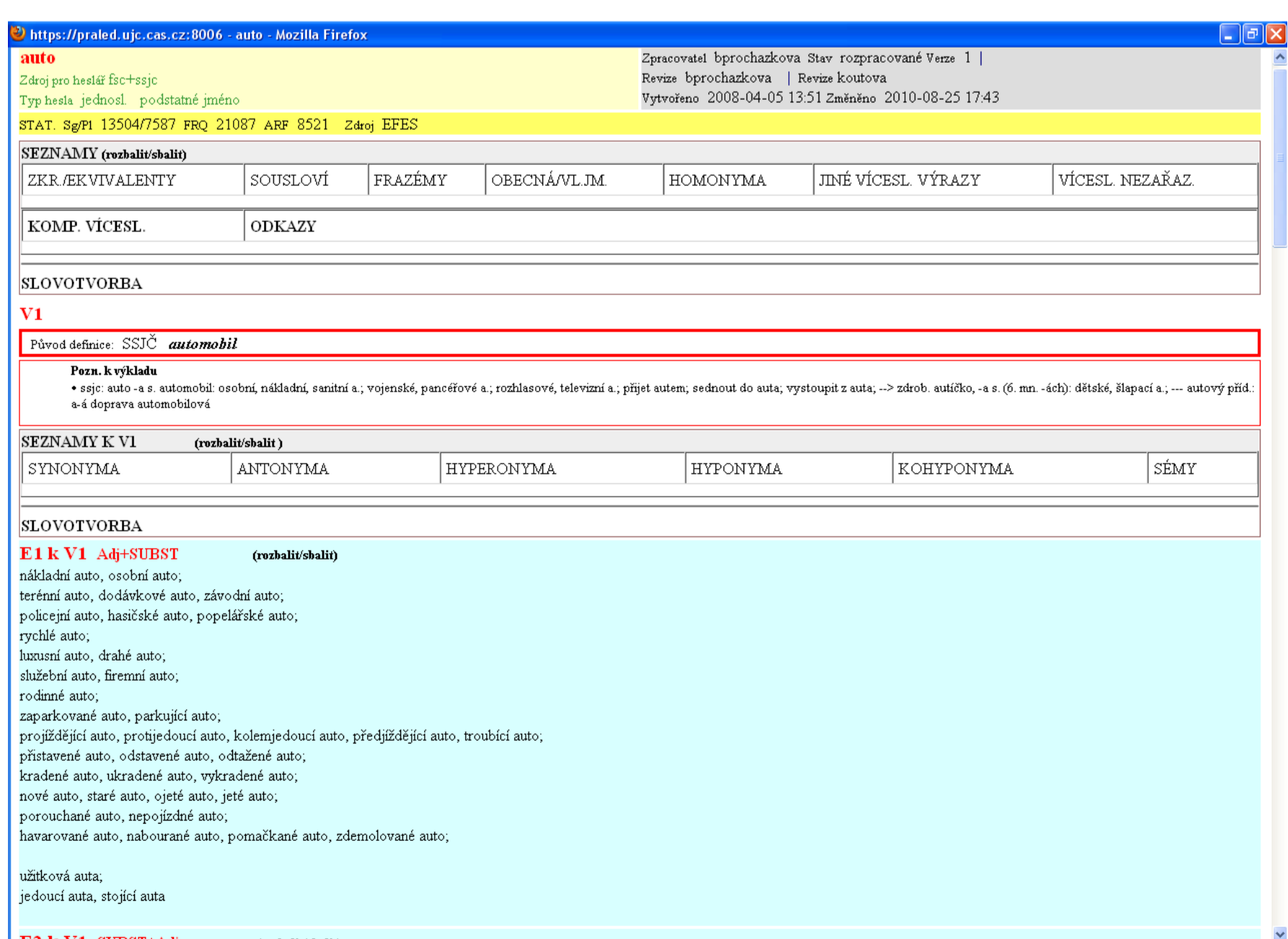
Exemplifikace:

K příkladům významům rovnou přiřazujeme lexikograficky relevantní doklady představující celou šíři spojitelnosti heslových slov, se kterými se pak bude podle potřeby pracovat dál (pro různé úzky). V příkladech číslujeme databázové hesla uvádíme tzv. **citátové doklady**, čili autentické větné kontexty s uvedením data, zdroje a typu textu i tzv. **upravené doklady**, které představují minimální syntagmata až vybrané části vět, dokládající celou oblast významového užití heslových slov a jejich jednotlivých významů (blíže viz v příspěvku M. Koutové).

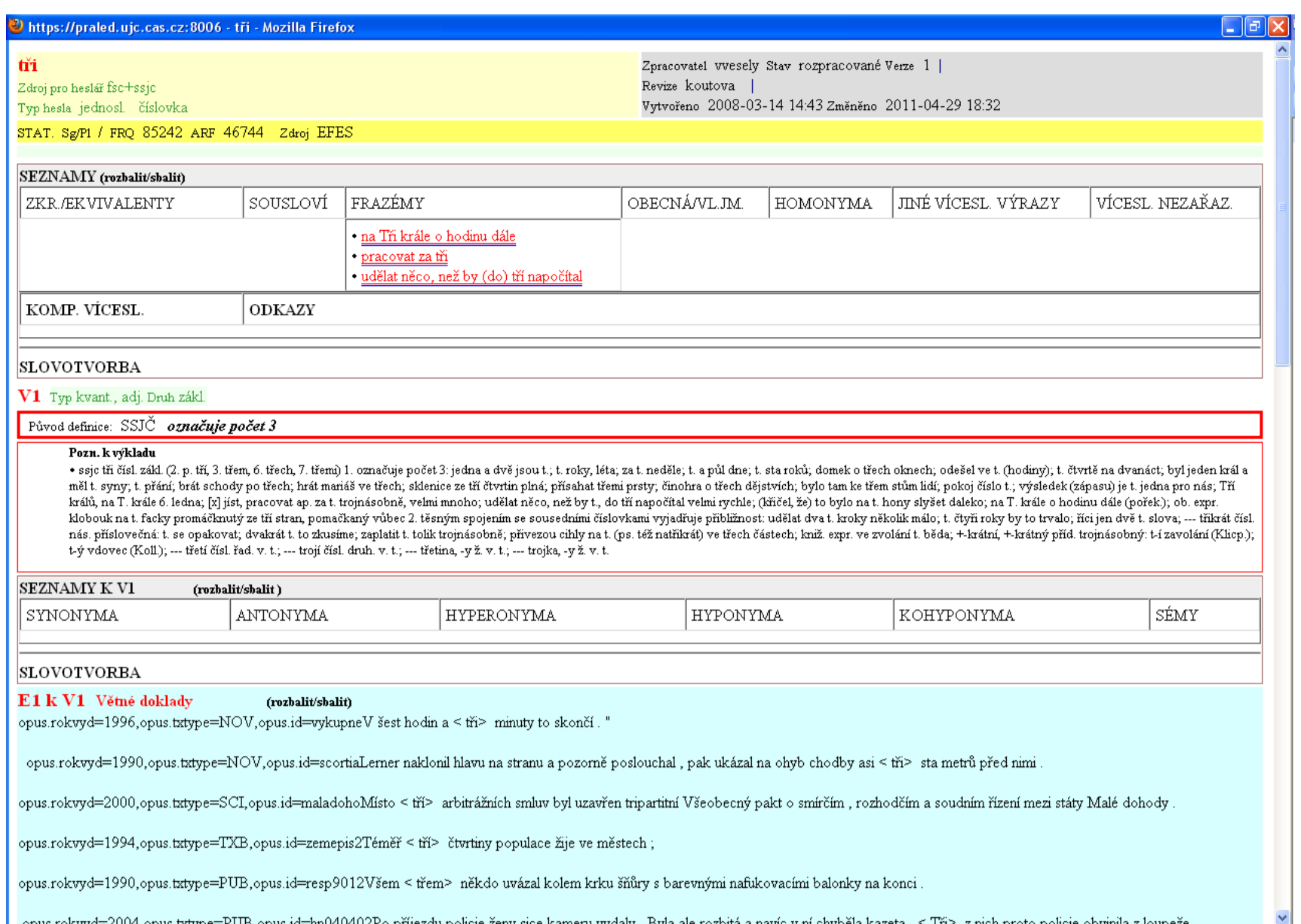
4 UKÁZKOVÁ HESLA Z RŮZNÝCH SLOVNÍCH DRUHŮ



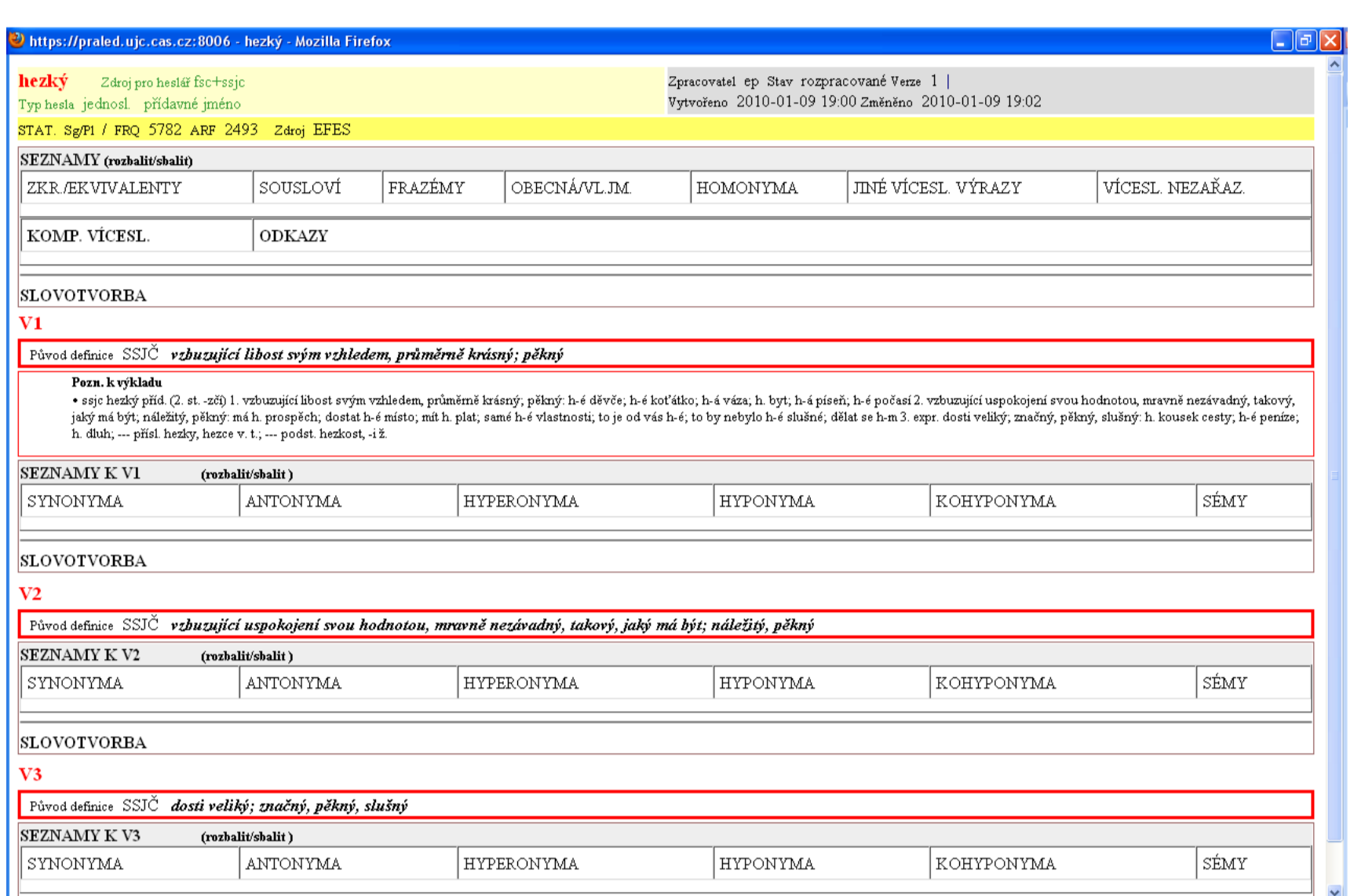
obr. 7: heslo *auto*



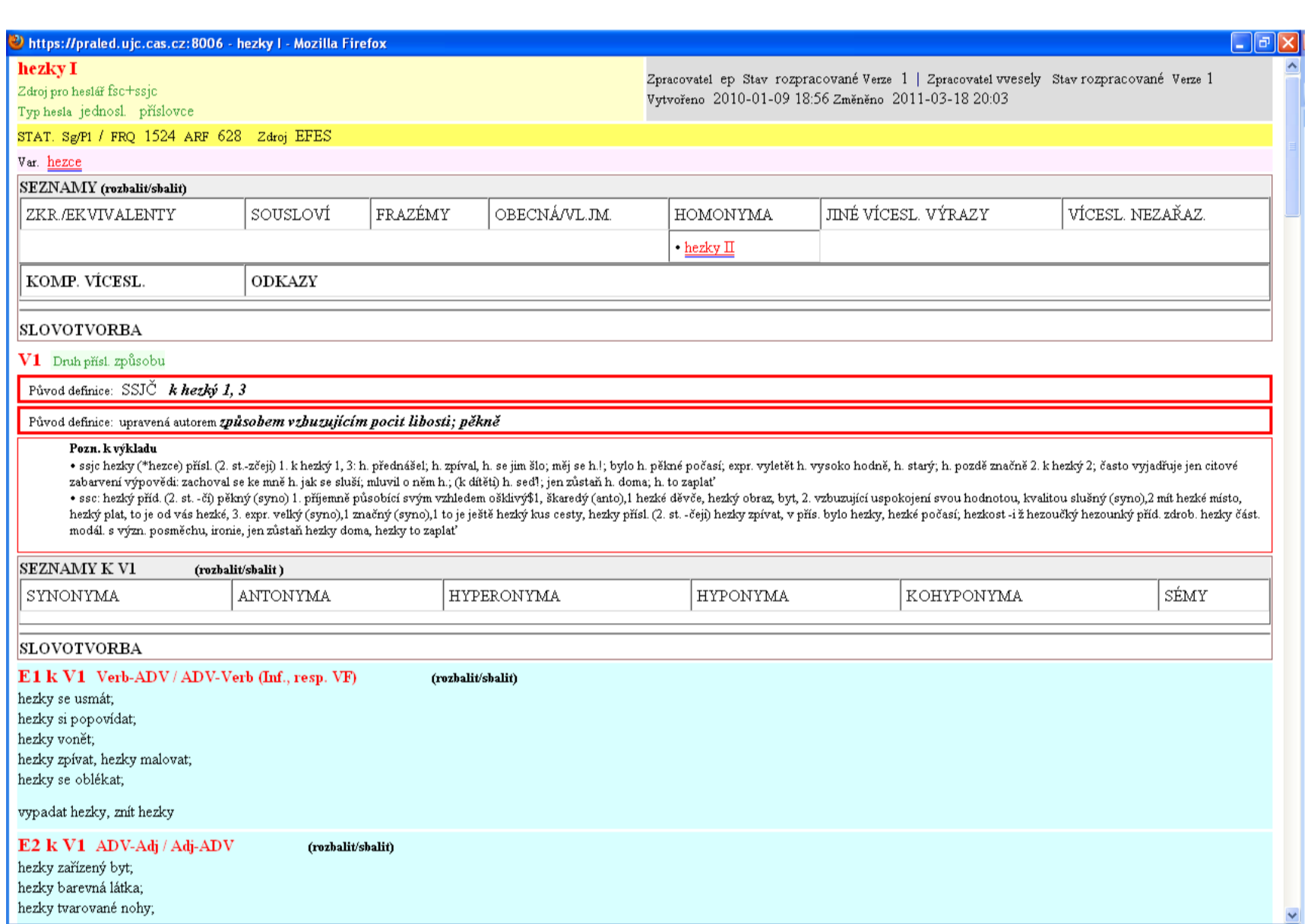
obr. 8: heslo *automobil*



obr. 9: heslo *řiti*



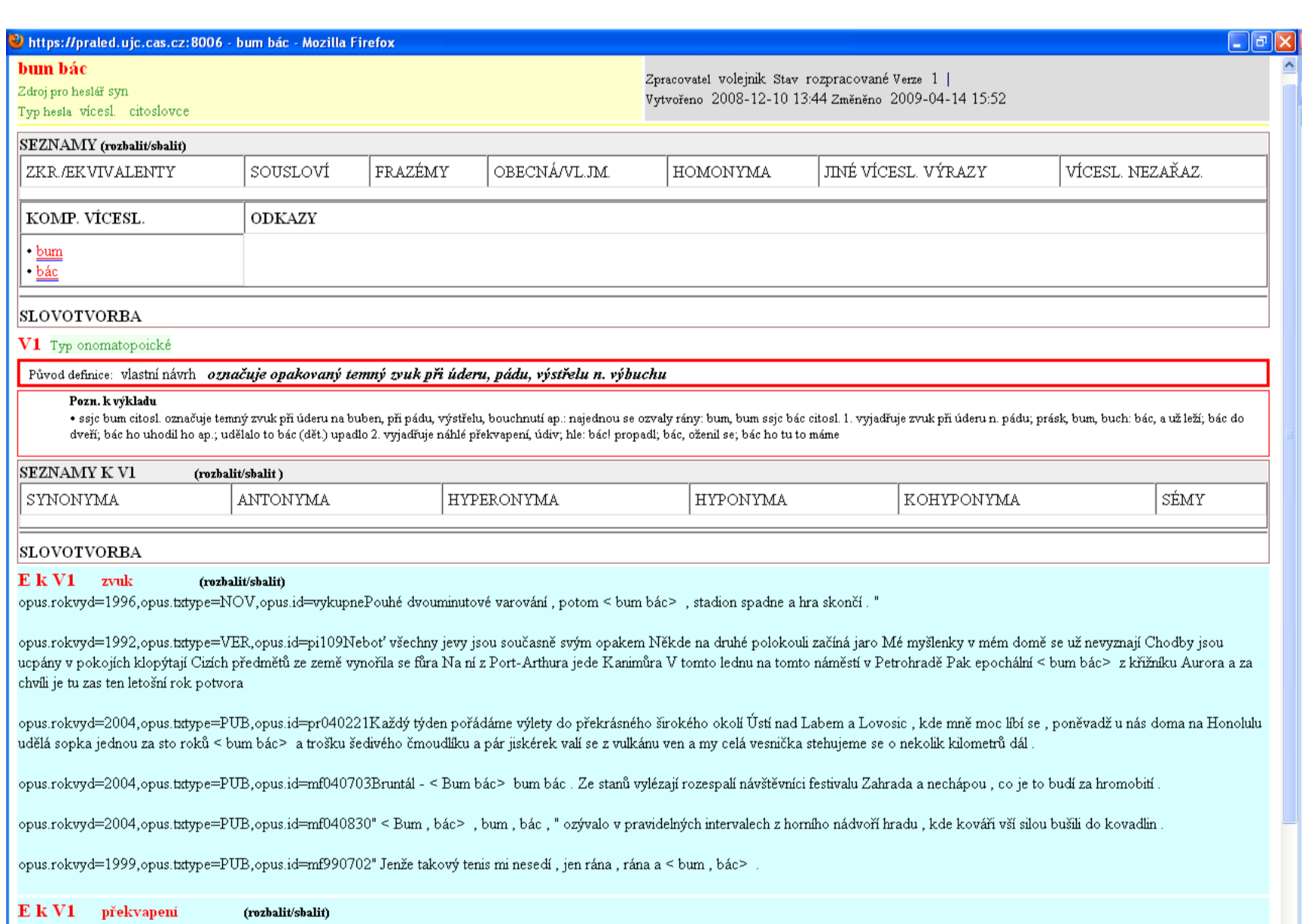
obr. 10: heslo *hezký* (jen založeno)



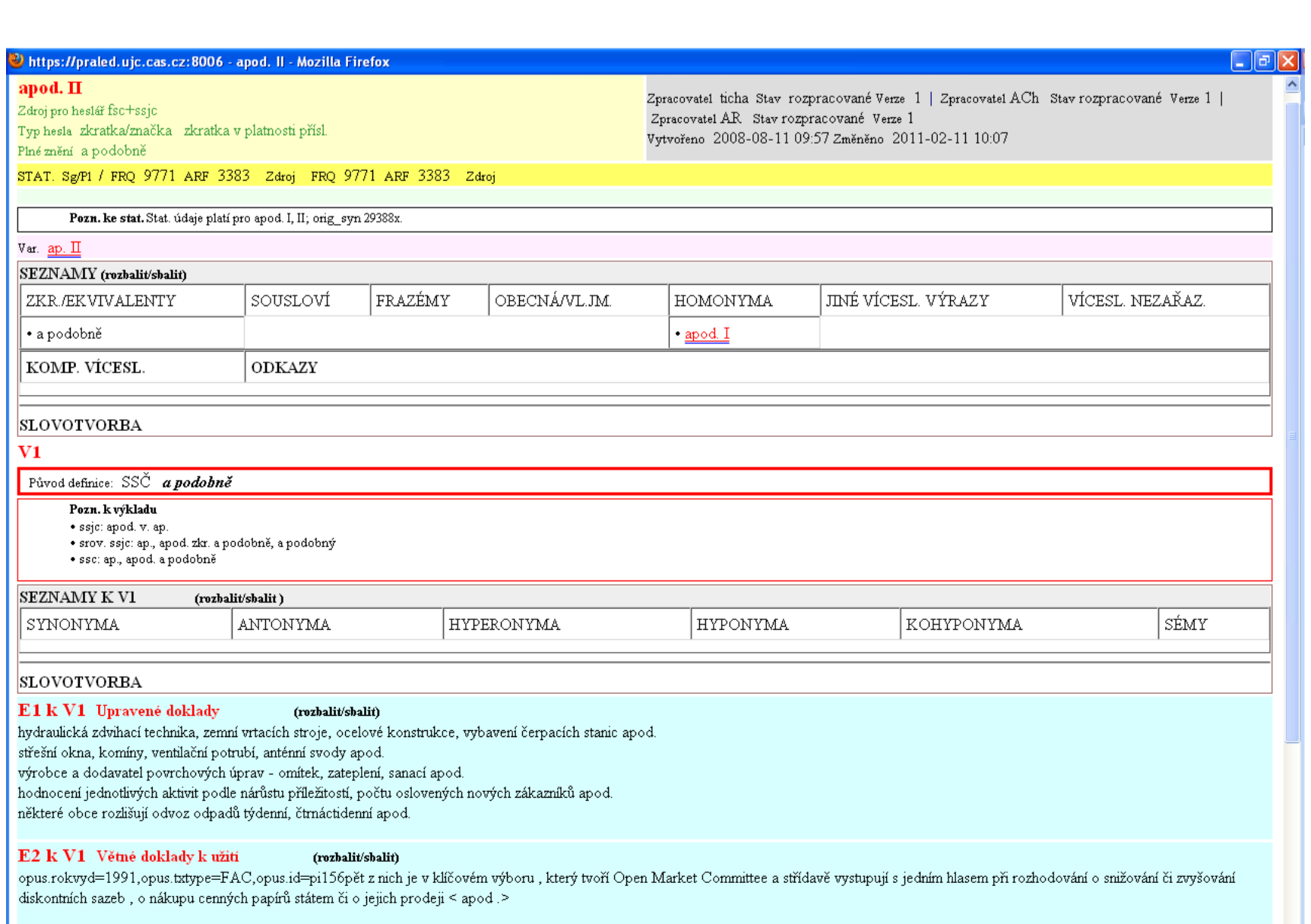
obr. 11: heslo *hezký I*



obr. 12: heslo *bác*



obr. 13: heslo *bum bác*



obr. 14: heslo *apod. II*