

**JIHOČESKÁ UNIVERZITA V ČESKÝCH
BUDĚJOVICÍCH**

**PEDAGOGICKÁ FAKULTA
KATEDRA ROMANISTIKY**



Lucie HRACHOVÁ

**Francouzská a česká terminologie v evropských
legislativních textech zaměřených na oblast: „Ochrana
zdraví“**

Vedoucí bakalářské práce: Mgr. Jan Radimský, Ph.D.

České Budějovice 2006

Bakalářská práce

Jihočeská univerzita v Českých Budějovicích

Pedagogická fakulta

Katedra romanistiky

Autor: Lucie Hrachová

Název: Francouzská a česká terminologie v evropských legislativních textech zaměřených na oblast: „Ochrana zdraví“

Vedoucí práce: Mgr. Jan Radimský, Ph.D.

Obor studia: Aplikovaná filologie – Anglický jazyk a francouzský jazyk pro administrativu Evropské Unie

České Budějovice 2006

Prohlášení

Tímto prohlašuji, že jsem bakalářskou práci zpracovala samostatně a uvedla v ní veškerou literaturu a ostatní zdroje, které jsem použila.

V Českých Budějovicích dne 25. 04. 2006

Hrachová' Lucie
.....

Poděkování

Velice ráda bych poděkovala Mgr. Janu Radimskému, Ph.D. za jeho trpělivost, ochotu a velmi cenné připomínky při vedení mé bakalářské práce. Můj velký dík patří i Janu Matějkovi za jeho rady týkající se informačních technologií a Ing. Jiřímu Ševčíkovi za jeho podporu a poznámky ke stylistické stránce mé práce. V neposlední řadě bych ráda vyjádřila vděk své rodině, jež mě při mém studiu velice podporuje.

OBSAH

1. ÚVOD	7
2. KORPUSOVÁ LINGVISTIKA	8
2.1 ÚVOD	8
2.2 CO JE TO KORPUS.....	8
2.3 HISTORIE KORPUSOVÉ LINGVISTIKY	9
2.4 TYPY KORPUSŮ	10
2.5 ČESKÝ NÁRODNÍ KORPUS (ČNK)	13
2.5.1 Vznik ČNK.....	13
2.5.2 Struktura ČNK.....	15
2.5.3 Budoucnost ČNK.....	17
3. ANALÝZA TEXTŮ	19
3.1 ÚVOD	19
3.2 PROGRAM PARACONC.....	19
3.2.1 Úprava textů pro práci s programem ParaConc	20
3.2.2 Nahrání textů do programu ParaConc	21
3.2.3 Zarovnání textů v ParaConcu.....	22
3.2.4 Vyhledávání v paralelních textech	24
3.2.5 Frekvence slov.....	28
3.3 TEXTY Z OBLASTI „OCHRANA ZDRAVÍ“	30
3.4 KLÍČOVÁ SLOVA Z OBLASTI „OCHRANA ZDRAVÍ“	35
3.4.1 Klíčová slova a jejich kolokace přeložené chybně.....	36
3.4.2 Klíčová slova a jejich kolokace s nepřesným překladem.....	40
3.4.3 Klíčová slova a jejich kolokace s nesrovnalostmi ve francouzských textech.	41
3.4.4 Klíčová slova a jejich kolokace překládané více synonymy	44
3.4.5 Klíčová slova a jejich kolokace přeložené jednoznačně	54
3.4.5 Výsledky analýzy	59
4. ZÁVĚR	61
5. PŘÍLOHY	63
5. 1 SLOVNÍČEK TERMÍNŮ A JEJICH KOLOKACÍ	63

5.2 SMĚRNICE KOMISE UPRAVUJÍCÍ NÁZVY „ŘEPKA“ A „ŘEPICE“	69
PRAMENY A LITERATURA.....	76
POUŽITÉ ZKRATKY	78
RÉSUMÉ (EN FRANÇAIS).....	79
ANOTACE.....	80
ANNOTATION	81

1. ÚVOD

Cílem mé bakalářské práce je nastínit základní informace o poměrně nově vzniklém oboru v oblasti lingvistiky - tzv. Korpusová lingvistika a pomocí paralelního korpusu podrobněji prozkoumat překlady vybraných termínů z evropských legislativních textů daného sémantického pole, což je v mém případě „Ochrana zdraví“. Pokud se budou překlady lišit nebo budou obsahovat nějaké nesrovnalosti, pokusím se tyto jevy vysvětlit. Neopomenu popsat přesnější postup práce s vybranými texty, při níž jsem využívala moderních informačních technologií.

V první kapitole bych se chtěla věnovat již zmíněné Korpusové lingvistice. Položím si otázky: „Co je to korpus?; K čemu nám slouží?; Jaké druhy korpusů rozlišujeme?“. Dále bych pak chtěla věnovat pozornost Českému národnímu korpusu – ČNK, jehož součástí je i paralelní korpus a nově vznikající projekt Intercorp, podle něhož vzniká i nový projekt paralelních korpusů na katedře romanistiky PF JČU v Českých Budějovicích vedený Mgr. Janem Radimským, Ph.D., jehož součástí jsou i mé zpracované texty.

Druhá kapitola se pak zabývá samotnou výzkumnou částí, kdy v prvé řadě popíši, jak je třeba postupovat, abychom v závěru mohli texty ve dvou jazycích porovnávat, důležité je i zmínit práci s programem ParaConc, bez něhož by paralelní korpus nemohl fungovat. Poté bych se chtěla soustředit na samotný průzkum překladů. Tato část bude obsahovat podrobný popis vybraných textů z oblasti „Ochrana zdraví“, tzn. číslo dokumentu, pod kterým vystupuje v úředním věstníku Evropské unie na stránkách <http://europa.eu.int/eur-lex/>, dále pak rok vydání, číslo řady úředního věstníku a číslo úředního věstníku, podle kterých je možné text na výše uvedené adrese vyhledat. Nebudou chybět ani přesné názvy textů jak v jazyce českém tak i francouzském a údaje o počtu normostran a počtu slov.

Přílohu bude tvořit slovníček z vybraných termínů z dané oblasti, na kterých budu provádět podrobnější průzkum homogenity (resp. nehomogenity) překladů.

2. KORPUSOVÁ LINGVISTIKA

2.1 Úvod

Korpusová lingvistika je poměrně mladý jazykový obor, který vznikl na počátku 60. let dvacátého století. Je třeba říci, že tento obor není novou teorií jazyka, jen využívá moderní informační technologie, které dříve neexistovaly, ke zkoumání textů a skrze ně dospívá k poznání obecnějších jazykových zákonitostí a pravidel. Jde vlastně o úzkou spolupráci lingvistů s jinými obory jako je matematika a počítačová věda. Existence počítačů také souvisí s tím, proč se u nás tento obor vyskytuje až v devadesátých letech minulého století. Informační technologie se u nás začaly rozvíjet až po pádu železné opony a teprve tehdy se čeští jazykovědci začali zajímat, jak to vlastně na západě s tvorbou novodobých národních korpusů funguje. (SCHMIEDTOVÁ, 1999).

2.2 Co je to korpus

Původem latinské slovo korpus (corpus, - oris) znamená 1. celek, soupis, sbírka 2. tělo, těleso, tělísko, postava 3. rozsáhlý soubor elektronicky uložených jazykových textů nebo jejich částí určený k vědeckému výzkumu jazyka. (Nový akademický slovník cizích slov A-Ž, 2005 : 141; 446).

Textový korpus je vlastně velká jazyková databanka, velmi rozsáhlý a složitý systém elektronicky zpracovaných textů z různých oborů a slouží k jazykovému výzkumu velmi moderní formou a to za pomoci počítačové techniky. Práci s korpusem umožňuje speciální vyhledávací program, pomocí něhož je možné vyhledávat slova a

slovní spojení v kontextu a zjistit jejich frekvenci výskytu. V dosavadním vývoji lingvistiky je tato forma výzkumu významným převratem. Počítače jsou nejen přesnější než lidé, ale i dostatečně rychlé k tomu, aby mohly zpracovat velké množství dat, což by bylo ruční analýze tak velkého počtu textů nemožné. Svým přesným měřením a rychlým vyhodnocováním dat jsou vlastně nenahraditelné.

„Korpus je tedy největší informační základna vůbec. Slouží jako zdroj materiálu nejen lingvistům, ale i širokému spektru zájemců z jiných oborů k mnohostrannému poznání jazyka, zákonitostí lidského myšlení a kultury.“ (ŠULC, 1999 : 11).

Klíčovým slovem pro jazykové korpusy je tzv. „reprezentativnost“. Před vznikem jakéhokoliv korpusu si musí tvůrci stanovit jeho cíl. Zda to bude korpus, který bude zkoumat jazyk například jednoho autora, nebo jednoho oboru a nebo zda bude zkoumat jazyk jako jeden celek nebo paralelní výzkum několika jazyků. Když si stanoví tento cíl, pak velice záleží na výběru vhodných textů, které budou pestré a budou co nejvíce reprezentovat daný cíl, na který je daný korpus zaměřen. Mohou to být novinové články, krásná literatura, odborné publikace aj. Je třeba podotknout, že pro studium jevů celotextové povahy je zapotřebí velkého množství textů.

2.3 Historie korpusové lingvistiky

Dřívější generace lingvistů již takovým způsobem, jakým funguje novodobá korpusová lingvistika, pracovaly i bez informačních technologií. Bylo to ale daleko náročnější a navíc si nebyli nikdy jisti, zda pracují s dostatečnou materiálovou základnou a zda jsou jejich pozorování a závěry přesné. Jednalo se vlastně o takovou „mravenčí práci“, kdy lingvisté a jejich pomocníci vypisovali z různých materiálů slova na lístky a k nim přiřazovali kontexty. Byl u nich uveden i přesný zdroj, autor, vydavatel, stránka apod. Tyto ruční výpisky se odborně nazývaly *excerpta*. (BLATNÁ, 1997 : 670). Ta se poté ukládala do osobních archivů nebo archivů vědeckých ústavů a z nich v minulosti vznikaly mluvnice, slovníky a jazykové učebnice. Takovým způsobem vznikl v minulosti například tezaurus Jana Amose Komenského nebo Jungmannův česko-německý slovník, jak uvádí Renata Blatná (1997 : 17). Je ale

zřejmé, že takovým způsobem by nemohl vzniknout textový korpus o obsahu 100 a více milionů slovních tvarů, jak je tomu dnes s využitím počítačové techniky.

První obecný (nespecializovaný) jazykový korpus vznikl v roce 1961, jeho tvůrci byli Nelson Francis a Henry Kučera a nazývá se Brown Corpus (BC)- „Brown University Standard Corpus of Present-Day Edited American English“. Tento korpus patří mezi korpusy s malým obsahem, má zhruba jeden milion textových slov. Brown Corpus byl prvním elektronickým korpusem vytvořeným v první řadě pro lingvistické účely. Bylo ale zřejmé, že malé korpusy mohou být využity jen na některé oblasti lingvistiky a že pro daleko přesnější informace o textových slovech je zapotřebí vytvořit korpusy daleko větší. (ŠULC, 1999 : 30).

Průkopníkem v tvorbě velkokapacitních korpusů byla Velká Británie, kde v 80. a na začátku 90. let vznikl projekt COBUILD (Collins Birmingham University International Language Database) Corpus vedený J. Sinclairem. V roce 1990 byla stávající databáze rozšířena a vznikl obrovský korpus Bank of English, který obsahuje v dnešní době více než 500 milionů slovních tvarů. V těchto letech vzniká také British National Corpus (BNC) o rozsahu asi 100 milionů slov, který má i významnou složku mluvenou a je to korpus, který se dále nerozšiřuje, je to tzv. uzavřený korpus. (ŠULC, 1999 : 35-37). Tento projekt je pro svou velikost a složení považován za vzor pro ostatní národní korpusy. I jiné evropské země mají své národní korpusy. Známý je například francouzský korpus Frantext nebo korpusy německé. Obecně se ale dá konstatovat, že v dnešní době snad není evropská země, která by svůj národní korpus neměla nebo by na něm nepracovala. (<http://knihovna.nkp.cz>).

2.4 Typy korpusů

Jazykové korpusy lze rozdělit do několika skupin na základě různých kritérií.

Prvním kritériem je časové rozmezí, ze kterého pocházejí texty daného korpusu a jedná se o korpusy synchronní a diachronní. (ŠULC, 1999 : 12).

Synchronní korpus (synchronic corpora) se vyskytuje nejvíce a zachycuje současný jazyk. Analýza těchto korpusů je nejpotřebnější a texty tvořící tento korpus jsou zároveň nejdostupnější (především v elektronické podobě). Skutečná časová hranice, podle níž by se dalo určit, kdy se jedná o korpus synchronní neexistuje. Obecně by se dalo říci, že se jedná o texty z několika posledních desetiletí s ohledem na společenské či jiné vnější přeměny. Synchronní korpusy jsou vysoce využitelné zejména z důvodu velkého množství textů a informací v nich obsažených stejně jako kvůli nejrůznějším poučením o jazykových i nejazykových jevech o jejich výskytu i užívání v přirozeném jazyce.

Diachronní korpus (diachronic corpora) naopak zachycuje texty mnohem starší. Pokrývá několik vývojových stádií daného jazyka nebo celý jeho vývoj. Oproti synchronním korpusům, které mají k dispozici neomezené množství textů ke zkoumání, je korpus diachronní daleko menší, má rozsah obvykle kolem 2000-5000 slov. Je tvořen ze starých dochovaných textů, jejichž počet je obzvláště ve starších vývojových fázích jazyka dosti omezený. Jedná se především o náboženské texty, legendy nebo texty veršované, jež mohou být použity v korpusu jen v elektronické podobě, takže je třeba je buď naskenovat a nebo využít méně častou metodu, což je přímý přepis písáčkou.

Dalším kritériem dělení je cíl neboli záměr korpusu. V tomto případě bychom korpusy rozdělili na obecné (všeobecné) nebo specializované. (ŠULC, 1999 : 12).

Korpusy obecné (general corpora) jsou takové, které nejsou vytvářeny pro nějaký určitý cíl nebo pro specifický jazykový výzkum. Snaží se zachytit jazyk v co největší šíři a snaží se reprezentovat jazyk jako celek. Jsou využívány zejména pro tvorbu slovníků, gramatik a jiných jazykových příruček.

Opakem jsou **korpusy specializované** (specialized corpora), které jsou zaměřeny na výzkum určité, specializované oblasti. Tento cíl samozřejmě korpus ovlivňuje, zvláště pak texty, ze kterých je korpus tvořen. Tento typ je využíván například při výzkumu různých nářečí nebo se může jednat o autorský korpus (např. korpus díla Bohumila Hrabala apod.). (<http://ucnk.ff.cuni.cz/korpus/korpus.html>).

Pokud bychom chtěli korpusy rozdělit z hlediska zachycení běžné každodenní komunikace, pak se jedná o korpus mluvený (spoken corpora), jehož opakem je korpus psaný (written corpora).

Mluvený korpus se bohužel vyskytuje velmi málo a jeho obsah také není příliš velký. V případě mluveného korpusu mluvíme samozřejmě pouze o synchronním, jelikož o diachronním se nedá ani uvažovat. Tvorba mluveného korpusu je ale zatím velmi náročná a drahá. Oproti *psanému korpusu*, který má k dispozici texty většinou již v elektronické podobě, je potřeba mluvený jazyk, zaznamenaný například na magnetofonové pásce, nejprve převést do počítačové podoby a teprve potom ho lingvisticky zpracovat. Problémem také zůstává, kde takové nahrávky vlastně získávat. Pokud bychom využili například veřejných projevů získaných z médií, chyběla by jim jistá spontánnost či neformálnost. Jednalo by se většinou o projevy předem připravené či přímo čtené. V rámci mluveného korpusu se objevují snahy vybudovat nářeční korpusy, kde ale nastávají komplikace s reprezentativností vzorků, které by zcela jistě nemohly být aplikovány obecně pro všechny sociální, věkové či kulturní skupiny. (KOCEK, KOPŘIVOVÁ, KUČERA, 2000 : 8).

Reprezentativnost korpusů je dalším rysem, podle kterého bychom mohli korpusy dělit. V tomto případě se jedná o *korpusy obsahující všechny možné texty*, které jsou předmětem jazykového výzkumu. Pokud vezmeme jako příklad výzkum díla již zmíněného Bohumila Hrabala, pak by bylo v tomto případě zapotřebí získat všechna jeho díla a na nich pak pomocí korpusu aplikovat jazykový výzkum. V jiném případě se jedná o *korpusy založené pouze na vzorcích*, které ale musí být vybírány pečlivě s ohledem právě na reprezentativnost jazyka jako celku. Tyto korpusy pak o jazyku podávají informace obecné. (ŠULC, 1999 : 13).

Podle množství obsažených jazyků můžeme korpusy rozdělit na *korpusy jednoho jazyka*, což je například právě Český národní korpus, nebo na *paralelní korpusy* (parallel corpora), kde se vyskytují texty ve dvou či více jazycích (většinou originál a překlad). (ŠULC, 1999 : 12). Pomocí zrovnávacích (párovacích) programů se oba soubory dají pozorovat souběžně a lze v nich vyhledávat slova, slovní spojení i jejich kolokace. Tyto korpusy jsou v poslední době velmi aktuální, neboť nabízí netradiční pozorování různých překladů a zároveň obohacují obyčejné dvoujazyčné slovníky o různorodější cizojazyčné ekvivalenty používané dobrými překladateli. Velmi významné jsou také pro výuku studentů, překladatelů apod. (KOCEK, KOPŘIVOVÁ, KUČERA, 2000 : 9).

Mimo výše jmenované druhy korpusů vznikají ještě *korpusy studijní* (learners' corpora), obsahující převážně texty psané studenty cizích jazyků, nebo *cvičné a testovací korpusy*, které mají čistě technickou povahu. Pomocí těchto korpusů se například trénují a vylepšují různé lingvistické hypotézy. Existuje i dlouhá řada *čistě tématických korpusů* vytvářených pro jeden obor či odvětví. (<http://knihovna.nkp.cz>).

2.5 Český národní korpus (ČNK)

Český národní korpus je rozsáhlý projekt obsahující několik jednotlivých korpusů (viz 2.5.2), který byl vytvořen pro lepší monitorování a výzkumy českého jazyka a je určen široké škále zájemců, nejen lingvistům. Dalo by se konstatovat, že je to dnes nejobsáhlejší forma jazykových informací a dat, které se dají díky počítačové technice využívat pro mnoho různých účelů. Jelikož je Český národní korpus nekomerční, mohou ho využívat lingvisté i nelingvisté, učitelé i studenti, badatelé či odborníci a v neposlední řadě zájemci o nový způsob využívání informací o českém jazyce nebo třeba jen proto, aby si obohatili svou slovní zásobu.

2.5.1 Vznik ČNK

Snahy o vytvoření korpusu se u nás objevovaly již před vznikem samotného ČNK a to v roce 1988, kdy pod záštitou Kybernetické společnosti vznikla *Iniciativní skupina pro přípravu počítačových korpusů textů a slovníků*. Tato skupina řídila nejen aktivity uvnitř našeho státu, ale snažili se i o spolupráci se zahraničím, které bylo v tvorbě jazykových korpusů daleko před námi. V roce 1990 se konalo sympóziu sdružení Language Industries (sdružení vytvořeno v roce 1986 zeměmi Evropského společenství s cílem zachovat identitu a dědictví všech evropských jazyků prostřednictvím moderního zpracovávání informací – pomocí počítačové technologie), kde bylo jasně oznámeno, že pro budoucí spolupráci s Evropským společenstvím

popřípadě pro získání podpory ze strany Společenství je zapotřebí, aby vznikly kvalitní národní jazykové projekty. (ŠULC, 1999 : 45).

Po pádu komunismu vznikla v roce 1991 významná iniciační skupina *Počítačový fond češtiny* složená z lingvistů a jiných odborníků především z oblasti matematiky. Účastníci této skupiny se snažili především vytvořit databázi textů, jež by mohly být zkoumané počítačem a pomocí ní pak zajistit lepší a kvalitnější tvorbu slovníků, gramatik a jiných jazykových příruček. V této skupině se postupem času začala plodit myšlenka o vytvoření projektu, který by měl celonárodní charakter. (ŠULC, 1999 : 46). První jednání vedené prof. Františkem Čermákem¹, tvůrcem výše zmíněného projektu, s vedením Ústavu pro jazyk český skončilo pro jeho neochotu neúspěchem. Mnohem úspěšnější bylo jednání s Filozofickou fakultou Univerzity Karlovy ve vedení s doc. Františkem Vrhelem, na jehož základě vznikl 9. září 1994 ÚČNK – *Ústav Českého národního korpusu*. Jeho pravidelná práce začala 1. října 1996 a jeho rozvoj byl podpořen několika granty a sponzory. (SCHMIEDTOVÁ, 1999).

Na tvorbě Českého národního korpusu se s ÚČNK podílejí i jiná významná pracoviště jako:

Fakulta informatiky MU Brno,

Ústav formální a aplikované lingvistiky MFF UK Praha (nástroje na morfologickou analýzu a značkování korpusů).

Ústav teoretické a počítačové lingvistiky - FF UK Praha

Ústav bohemistických studií - FF UK Praha

Katedra českého jazyka a teorie komunikace- FF UK Praha

Ústav pro jazyk český Akademie věd České republiky Praha

Ústav pro českou literaturu AV ČR

Ústav českého jazyka a slovanské jazykovědy FF MU Brno

Pedagogická fakulta MU Brno

Elektrotechnická fakulta ČVUT, katedra počítačů, Praha

(<http://ucnk.ff.cuni.cz/>), [cit. 8. 4. 2006].

¹ Prof. PhDr. František Čermák, DrSc., (*1940) vystudoval Filozofickou fakultu UK v Praze. Na této fakultě se v Ústavu Českého národního korpusu, který založil a jehož je ředitelem, zabývá zejména lexikologií a frazeologií. Obecně se zabývá lingvistikou a bohemistikou. V poslední době publikoval (spolu s I. Klímovou a V. Petkevičem) Studie z korpusové lingvistiky (Karolinum, Praha 2000). Je členem Učené společnosti ČR. (<http://www.ucnk.cz/>) [cit. 8. 4. 2006].

2.5.2 Struktura ČNK

Podobně jako jsem uvedla dělení jazykových korpusů obecně bychom mohli rozdělit i ČNK. Ústav českého národního korpusu se zabývá několika projekty a byl vybudován jak korpus synchronní tak diachronní, psaný i mluvený a nechybí ani projekt zabývající se korpusem paralelním. Jak je ČNK rozdělen a jaké jsou přesné názvy současných projektů lze vypožorovat z následující tabulky.

Tabulka č. 1: *Struktura ČNK*

Český národní korpus		
Synchronní část		Diachronní část
<p>Banka synchronní češtiny představuje současné české texty upravené pro vyhledávání korpusovým manažerem. V současné době technické prostředky ani různorodost zpracování sbíraných dat neumožňují pracovat se všemi složkami najednou, proto jednotlivé části vznikají jako samostatné korpusy:</p>		<p>Banka diachronní češtiny se skládá ze staročeských textů zařazených do: banky transkribovaných textů (2 miliony textových slov), banky transliterovaných textů (asi 100 tisíc textových slov) a banky nářečních textů (asi 200 tisíc textových slov).</p>
Psané korpusy	Mluvené korpusy	Diachronní korpus
<ul style="list-style-type: none"> • SYN2005 • SYN2000 • FSC2000 • PUBLIC • SYNEK • LITERA • ORWELL 	<ul style="list-style-type: none"> • Pražský mluvený korpus • Brněnský mluvený korpus 	<p>Diachronní korpus obsahuje výběr staročeských textů od prvních dochovaných záznamů (13. století) do doby pokryté synchronním korpusem.</p> <ul style="list-style-type: none"> • DIAKORP (v roce 2005 byla veřejnosti zpřístupněna první jeho část o rozsahu 700 000 slovních tvarů)
<p>Paralelní korpusy</p> <p>projekt InterCorp</p>		

Pokud bychom se chtěli na jednotlivé korpusy podívat podrobněji, tak největší ze všech jsou dva korpusy psané a to *SYN2000* a *SYN2005*. Jde o korpusy, které mají přes 100 milionů slov, tedy dohromady 200 milionů. U každého slova (tedy jeho výskytu v textu) jsou uvedeny mezinárodně kompatibilními značkami informace o jeho zdroji (tedy z jakého původního textu slovo pochází), o jeho morfologických vlastnostech a stejně tak je zde uvedena i jeho lemma (tedy základní tvar tohoto slova).

„Samotné texty v *SYN2000* mají rozsah 1–2 gigabytů a jejich prosté manuální prohlížení přesahuje lidské možnosti. Toto množství si lze při průměrné podobě a velikosti tištěných knižních stránek a při průměrné knize o 250 stranách tištěných na tenkém papíru představit jako 10 zaplněných metrů knihovny. Pouhé čtení celého synchronního korpusu by při poměrně rychlém tempu (150 slov za minutu, 8 hodin denně a 365 dní ročně) zabralo přes 4 roky.“ (KOCEK, KOPŘIVOVÁ, KUČERA, 2000 : 14). Avšak je třeba si uvědomit, že za pomoci moderní technologie je možné zjistit výsledek během chvilky, aniž bychom četli jedinou stranu všech textů.

Projekt *SYN2005* je nový a ještě nebyl lemmatizován a morfologicky označován. Od projektu *SYN2000* se liší jak stářím textů (*SYN2005* se skládá především z textů nových) a poté zastoupením druhů textů, které je znázorněno v následující tabulce.

Tabulka č. 2: Reprezentativnost korpusů

	SYN2005	SYN2000
beletrie	40 %	15 %
odborná literatura	27 %	25 %
publicistika	33 %	60 %

(<http://ucnk.ff.cuni.cz/syn2005.html>), [staženo 8. 4. 2006]

Pro zájemce, jež by si rádi vyzkoušeli práci s korpusem na vlastní kůži byl vytvořen korpus PUBLIC, který má stejnou reprezentativnost jako *SYN2000* (je vlastně jeho podsložkou), ale obsahuje jen 20 milionů slov. Je volně přístupný na <http://ucnk.ff.cuni.cz/>.

Pokud jde o mluvené korpusy, ÚČNK pracuje na dvou a to na *Pražském mluveném korpusu (PMK)* a *Brněnském mluveném korpusu (BMK)*. Jelikož je tvorba mluvených korpusů nákladná a nepříliš lehce proveditelná, počet slov samozřejmě nemůže být tak rozsáhlý jako u korpusů psaných. PMK má v současné době přes 700 000 slov a byl vytvořen z 304 magnetofonových nahrávek, které jsou anonymní a pocházejí z let 1988–1996. BMK je tvořen z 250 nahrávek a pocházejí z let 1994–1999. Z hlediska reprezentativnosti korpusu je třeba zmínit, že nahrávky byly pořizovány tak, aby v něm byla zastoupena čtyři kritéria a to: pohlaví mluvčího (muž–žena), věk (nižší–vyšší, hranice kolem 35 let), vzdělání (nižší–vyšší) a typ textů (formální–neformální). Oba tyto korpusy jsou zatím ve vývoji a v budoucnu lingvistům poslouží především v oblasti lexikonu a morfologie. (KOCEK, KOPŘIVOVÁ, KUČERA, 2000 : 15-16).

Cílem projektu Intercorp je vytvoření paralelních korpusů a to mezi češtinou a ostatními evropskými jazyky pomocí dat ČNK. Jedná se o jazyky jako je: angličtina, němčina, francouzština, španělština, italština, švédština, dánština apod. Zahrnuty budou i „neevropské“ jazyky jako je: arabština, ruština, japonština nebo čínština. Tento projekt je nový a v budoucnu přinese odborníkům i širokému spektru zájemců velké množství poznatků o cizích jazycích, jejich překladech, přinese zcela jistě zdokonalení překladových slovníků a jiných publikací o daných jazycích a v neposlední řadě bude velmi užitečný jak pro pedagogy tak i studenty. Podle harmonogramu projektu InterCorp se počítá s tím, že například v roce 2007 by měl tento korpus obsahovat minimálně 1 000 000 slovních tvarů. V roce 2010 už by měl být jeho rozsah přes 2 500 000 a o rok později by měl být tento korpus přístupný na internetu. (<https://trnka.ff.cuni.cz/ucnk/intercorp/>), [cit. 9. 4. 2006].

2.5.3 Budoucnost ČNK

Český národní korpus je projekt kontinuální, tedy projekt, který se bude neustále rozšiřovat. Předpokládaný rozsah korpusu v roce 2011 je jedna miliarda slovních tvarů. Tento korpus českého jazyka bude nenahraditelným prostředkem k vytváření nových, lepších slovníků, mluvnic, jazykových příruček nebo učebnic. Již v roce 2004 byl na základě korpusu SYN2000 vydán Frekvenční slovník češtiny. Bude pokračovat sběr

mluvených textů, tak aby mohly mluvené korpusy pokrýt všechny oblasti České republiky. Existuje i myšlenka vytvoření zvláštního korpusu, který by se zabýval pouze slovní zásobou z období totalitního režimu. Pro tvorbu terminologických slovníků bude zapotřebí vytvořit korpusy speciální zahrnující texty pouze určitého oboru. ÚČNK začal svoji práci včas, a proto je korpus český srovnatelný rozsahem a zpracováním korpusům anglickým. ČNK se řadí mezi nejvýznamnější počiny české kultury po roce 1990. (<http://knihovna.nkp.cz>).

3. ANALÝZA TEXTŮ

3.1 Úvod

Tato kapitola je výzkumnou částí mé bakalářské práce. Než přejdu k samotné analýze vybraných klíčových slov, ráda bych nastínila fungování programu ParaConc, pomocí něhož jsem daná slova zkoumala.

3.2 Program ParaConc

Tento software nabízí širokou škálu zkoumání překládaných textů, ať už je to analýza dvoujazyčné terminologie nebo třeba jen zkoumání různých překladů jednoho textu. Program ParaConc není vázán pouze na určité jazyky, ale můžeme v něm porovnávat i texty například anglicko-čínské nebo česko-francouzské aj. Dokáže paralelně porovnávat i více než jen dva texty, maximum jsou však texty čtyři. Díky němu můžeme vyhledávat zadané termíny a jejich překlady během pár vteřin, ačkoliv bez něj by to ve velkém množství textů bylo velmi obtížné, pokud by to vůbec bylo možné. Program ParaConc nabízí různé možnosti zkoumání, ať už je to samotný překlad termínu v různých kontextech, frekvence jeho výskytu či zobrazení jeho kolokací (slova, která se s daným termínem nejčastěji pojí). Tento software není určen jen pro lingvisty či překladatele, ale i pro ostatní zájemce, kteří si chtějí obohatit svou slovní zásobu nebo například prozkoumat překlady různých termínů. (<http://www.athel.com/para.html>), [cit. 12. 4. 2006].

3.2.1 Úprava textů pro práci s programem ParaConc

Abychom mohli s danými texty pracovat v programu ParaConc je třeba je „vyčistit“, tzn. odstranit veškeré tabulky, obrázky a poznámky pod čarou. Poté následuje úprava na stejný počet odstavců obou textů. Konce odstavců by se neměly objevovat uprostřed vět, stejně tak by se tam neměly objevovat ani tabulátory. Postup pro úpravu odstavců je následující:

1. Otevřeme soubor v programu MS Word
2. V horní liště klepneme na nabídku *Nástroje*, pak na *Možnosti*. V této nabídce vybereme *Zobrazení* a v části *Značky formátování* zaškrtneme políčka *Znaky tabulátorů* a *Konce odstavců*. Výběr potvrdíme tlačítkem *OK*.
3. Zkontrolujeme zda se symbol ¶ označující konce odstavce nebo symbol → označující tabulátory nevyskytuje uprostřed věty. Pokud ano, jednoduše ho v textu smažeme. Je vhodné odstranit i všechny znaky odstavce, které nejsou na koncích vět.
4. Počty odstavců daných textů si ověříme, když v menu *Nástroje* vybereme možnost *Počet slov*.

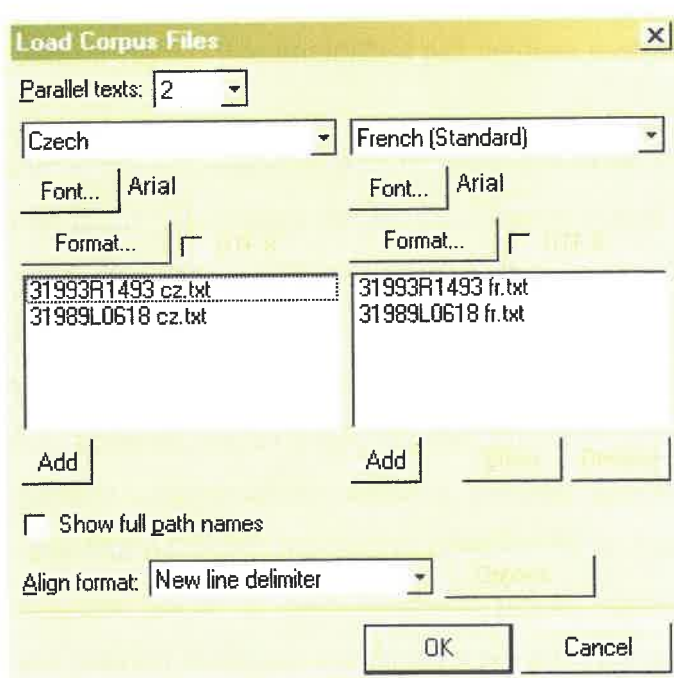
Až se nám podaří upravit texty na stejný počet odstavců, musíme je uložit do textového formátu (.txt), neboť ParaConc neumí pracovat s textem ve formátu DOC.

1. Soubor máme otevřený v programu MS Word. Klepneme na možnost *Soubor* a vybereme *Uložit jako*.
2. Jako *Typ souboru* vybereme „prostý text“ a dáme *Uložit*. Vyskočí okno s dotazem na typ kódování, které má být použito. V případě českých textů ponecháme možnost *Windows (výchozí)*, v případě textů francouzských zvolíme možnost *Jiné kódování* a z nabídky vlevo vybereme kódování *Západoevropské jazyky (Windows)*.
3. Potvrdíme tlačítkem *OK*. A postup opakujeme pro všechny texty.

3.2.2 Nahrání textů do programu ParaConc

Nyní by měly být texty připraveny pro nahrání do programu ParaConc. I když máme texty zkontrolované a upravené a počet odstavců by se měl shodovat, může se ještě objevit nějaká chyba, kterou ale lze upravit přímo v ParaConcu. Texty do ParaConcu nahrajeme následujícím způsobem:

1. Spustíme program ParaConc. V nabídce na horní liště vybereme *File* a zvolíme *Load Corpus File(s)* nebo použijeme klávesovou zkratku *Ctrl+L*. Nabízí se nám okno, kde je třeba zvolit texty, které mají být do ParaConcu nahrány. Texty zvolíme pomocí tlačítka *Add*, jak ukazuje obrázek č. 1.



Obrázek č. 1: Nahrávání textů

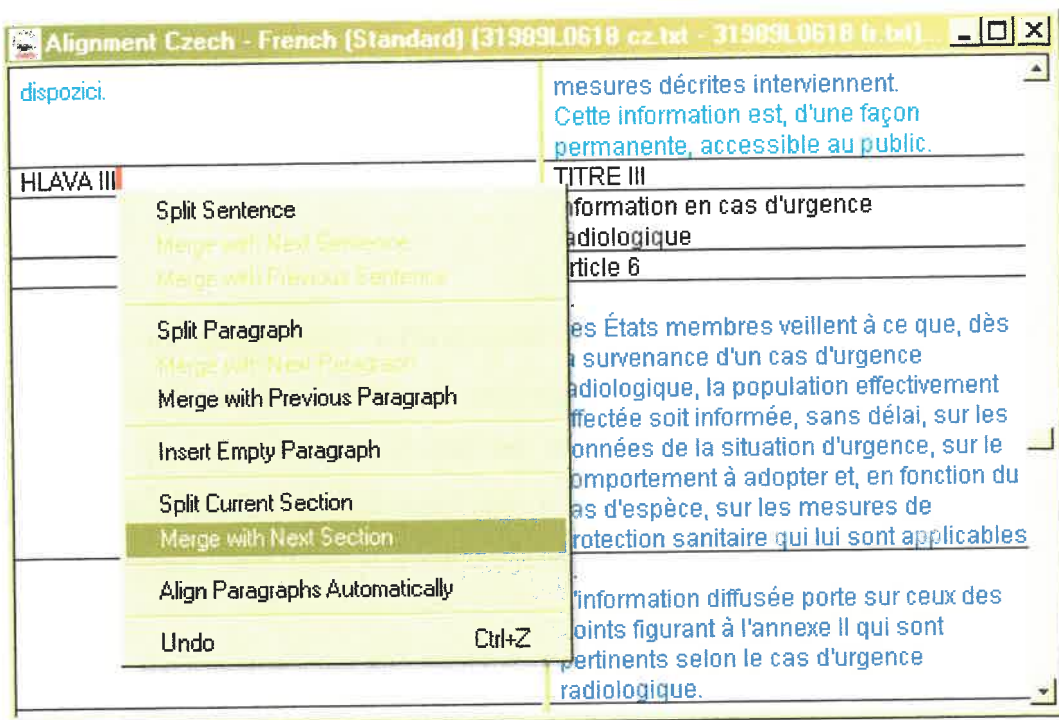
2. V horním okně s názvem *Parallel texts* ponecháme číslo 2, což znamená, že se chystáme nahrát pouze dva texty. Je možné zvolit nanejvýš texty čtyři. Z nabídky vybereme jazyky, které chceme v ParaConcu analyzovat. Důležité je zvolit vhodný *Font* písma, jelikož ne každý font je schopný zobrazit znaky charakteristické pro určitý jazyk. Já jsem zvolila font *Arial*, neboť je schopný bezchybně zobrazit jak texty české tak francouzské.

3.2.3 Zarovnání textů v ParaConcu

Nakonec je třeba v okně nahrávání textů zvolit *Align format* (formát zarovnání). Pokud máme texty upravené na stejný počet odstavců, pak zvolíme *New line delimiter*, program pak již nekontroluje, zda je počet odstavců správný a rovnou texty nahraje.

Jestliže zvolíme možnost *Not aligned*, program zkontroluje, zda je počet sekcí stejný a pokud ne, nahlásí chybu. V tomto hlášení pak zvolíme možnost *Fix*.

1. Když zvolíme ikonu *Fix*, objeví se nám na ploše dvě okna. První udává členění textu na sekce, druhá pak ukazuje členění na odstavce. Pokud se odstavec skládá z více vět, pak jsou věty rozlišeny barevně. Pro úpravu textů si vybereme tabulku se členěním na odstavce.
2. Nalezneme v tabulce místo, kde se nám paralelně zobrazené texty rozchází. Pak máme několik možností, jak texty upravit. Buď klikneme pravým tlačítkem myši na první písmeno „přebývajícího textu“ a z místní nabídky vybereme *Split Paragraph* (rozdělit odstavec) pro úpravu odstavce (nebo *Split Current Section*, pokud upravujeme sekce či *Split Sentence*, pokud upravujeme věty) nebo klikneme opět pravým tlačítkem tentokrát na poslední písmeno, za kterým text chybí a zvolíme možnost *Merge with...(spojit s...)* Opět můžeme tuto možnost aplikovat jak pro úpravu sekcí, tak i pro odstavce a věty. Jako ukázka poslouží následující obrázek:



Obrázek č. 2: Zarovnávání textů v ParaConcu

3. Po zarovnání tabulky zavřeme a postup opakujeme pro další texty. Když máme texty upravené, je třeba je uložit. V nabídce na horní liště vybereme *File* a *Save Workspace*. Příště už si pak nemusíme soubory znova načítat či zarovnávat, stačí jen otevřít ParaConcu a v nabídce *File* vybrat možnost *Open Workspace* (Ctrl+O).
4. Pokud chceme mít upravené i zdrojové texty, pak existuje samozřejmě možnost, že při zkoumání nevyhovujících odstavců je můžeme upravovat i v samotném textovém dokumentu. Potom máme jistotu, že kdykoliv si do ParaConcu nahrajeme texty znovu, nemusíme je již v programu zarovnávat.

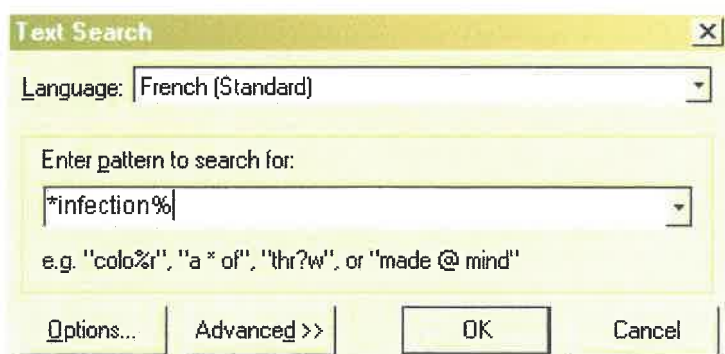
Pokud jsme na počátku nahrávání textů do ParaConcu zvolili jako *Align format* – *New line Delimiter*, program texty nekontroluje, ale rovnou je nahrává. I v takovém případě je vhodné si texty zkontrolovat a to tak, že v nabídce *File* vybereme možnost *View Corpus Alignment* (Ctrl+V), označíme texty, které chceme zkontrolovat a zvolíme *Alignment*. Opět se nám objeví dvě tabulky, jedna pro sekce a druhá pro odstavce. Na

několika místech v textu zkontrolujeme, zdali odstavce souhlasí a pokud ne, opakujeme postup v bodě 2 a 3.

3.2.4 Vyhledávání v paralelních textech

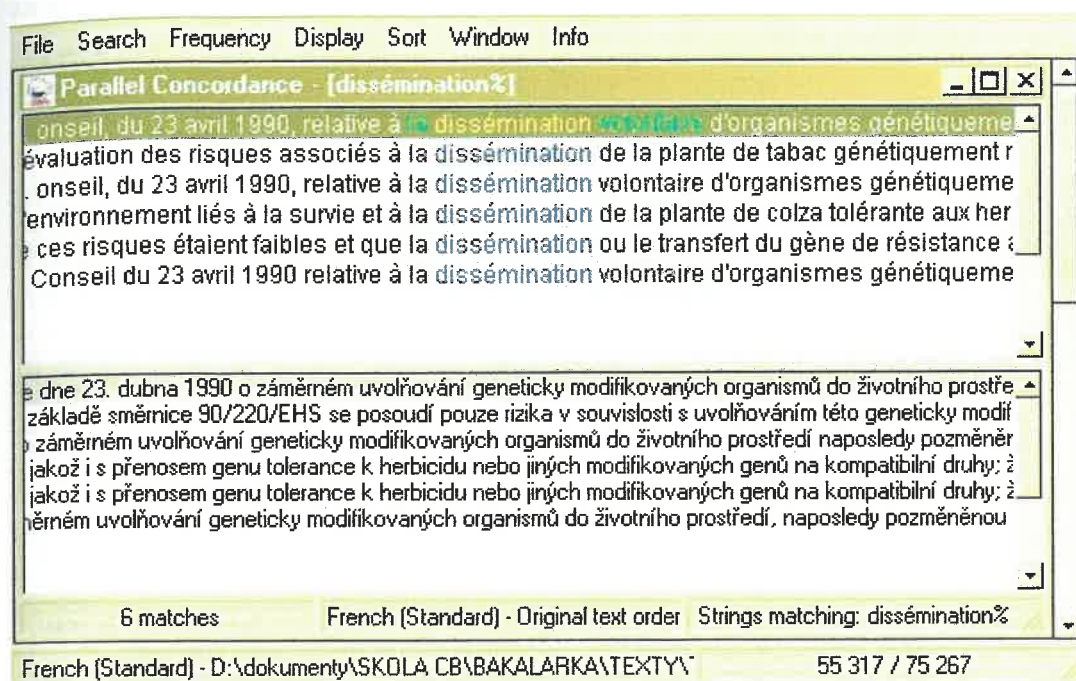
Nyní už máme texty v ParaConcu nahané a jsou upravené tak, že v nich můžeme vyhledávat slova, či slovní spojení, která nás zajímají. Tento program nám neposkytuje pouze možnost vyhledávání ale i informaci o celkovém počtu slov (v pravém dolním rohu) či frekvenci slov v textech, které jsou v něm nahané. Počet textů nahaných v ParaConcu je uveden v levém dolním rohu.

Termíny vyhledáváme tak, že v horní nabídce vybereme možnost *Search* v menu *Search* nebo použijeme klávesovou zkratku Ctrl+S. Je třeba dávat pozor již při zadávání hledaného termínu, jelikož ve francouzském jazyce jsou zrádná například substantiva začínající na samohlásku. Když zadáme do dotazového řádku takové slovo, počet vyhledaných termínů není korektní, neboť program opomene vyhledat tentýž termín, před kterým se ale nachází člen s apostrofem. V tomto případě takový termín pokládá za jiný. Je tudíž na místě zadat před takový termín symbol *. Důležité je také za substantiva umístit symbol %, aby program našel jak jeho singulár tak plurál. Jako ukázka poslouží tento obrázek:



Jakmile zadáme hledané slovo a odsouhlasíme tlačítkem *OK* program začne pracovat a v mžiku nám nabídne hledaný dotaz. V horní části okna se nám objeví tzv. „konkordance“, což je podle Renaty Blatné (1997 : 670) vlastně hledaný termín se svým textovým okolím. Ke zvýrazněným hledaným termínům jsou pak ve spodní části okna přiřazeny české věty, jež daný termín obsahují. Je to vlastně důsledek zarovnávacího procesu. Tzn., že když se hledaný francouzský termín vyskytuje v zarovnané části číslo 108, program k němu zkrátka přiřadí odpovídající část 108 v textech českých. Takto se pak uspořádají všechny ostatní segmenty hledaného slova. (<http://www.athel.com/para.html>). Když klikneme na jeden z vyhledaných termínů, v druhé části se zvýrazní úsek textu, kde se jeho ekvivalent v druhém jazyce vyskytuje. Pokud nás zajímá širší kontext termínu, můžeme na něj dvakrát kliknout a slovo v širším kontextu se objeví v novém okně.

Mimo hledané termíny nás budou jistě zajímat i jeho kolokace, což je smysluplné spojení dvou nebo více slov, zvláště v podobě víceslovného pojmenování (KOCEK, KOPŘIVOVÁ, KUČERA, 2000 : 50) nebo podle Nového akademického slovníku cizích slov (2005 : 418) zní definice termínu „kolokace“ následovně: „Kolokace je dvojice nebo i více slov, samostatných lexikálních jednotek obvykle spojovaných“. Aby se nám v programu slova, s kterými se náš daný termín pojí nejčastěji, zobrazila, je třeba v horní nabídce *Frequency*, zvolit *Frequency Options* a poté vybrat v *Collocate Span* možnost *1L-1R* (první zleva, první zprava). V tomto případě se nám zvýrazní každé více frekventované slovo, které se nachází nalevo nebo napravo od hledaného termínu, viz obrázek. 4.

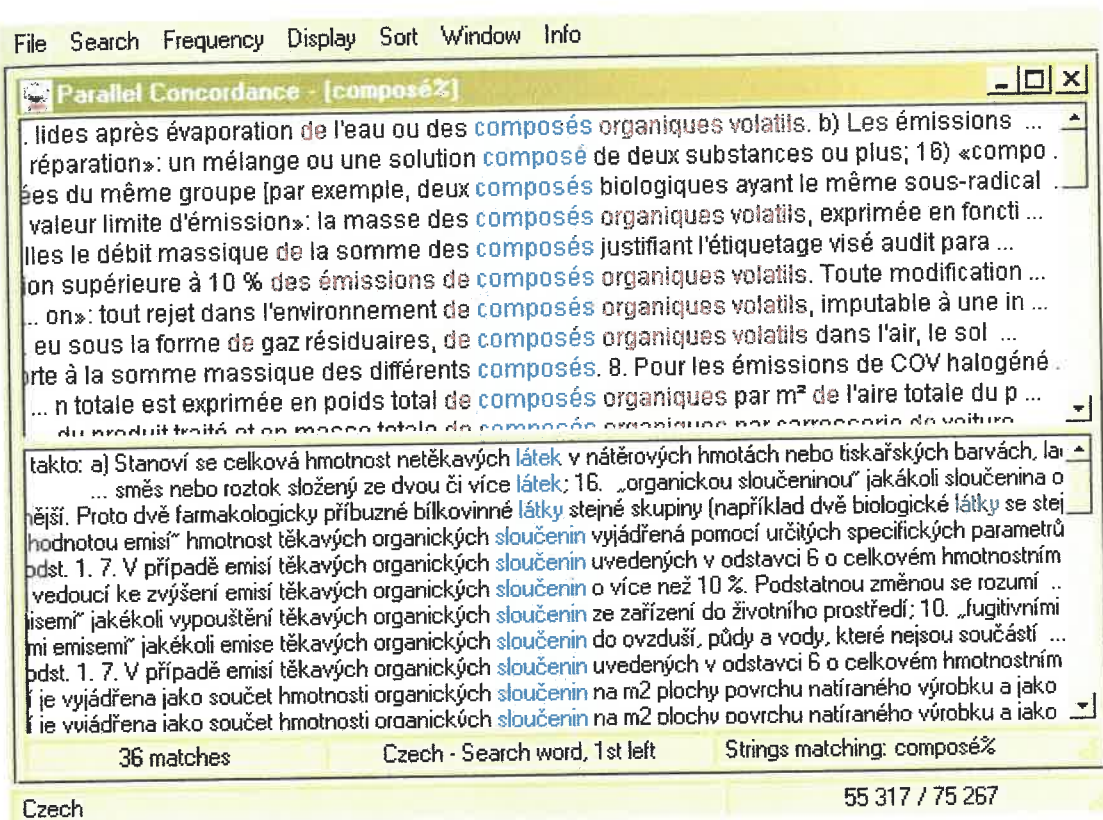


Obrázek č. 4: Vyhledaný termín a jeho kolokace

Nyní se ale setkáváme s problémem, jak najít ekvivalentní překlady ve spodní části okna, abychom nemuseli složitě pročítat všechny věty, které nám ParaConc nabídl. V tomto případě je třeba umístit kurzor na nějaké místo ve spodním okně, aby se stalo aktivním. Na nějakém místě v tomto okně klikneme pravým tlačítkem myši a v menu, které se nám nabídne vybereme *Search Query*, do dotazového řádku napíšeme české slovo (vhodný ekvivalent ke slovu francouzskému) a opět dáváme pozor na různé formy termínu, který zadáváme a použijeme symboly * a % či jiné (např. slova infekce, infekcím, ale i infekční či desinfekční) zadáme ve tvaru *infek*). Potvrdíme tlačítkem *OK* a ve spodním okně se nám zvýrazní vyhledaný dotaz. Nyní je již snadnější nalézt vhodný překlad termínů zvýrazněných v horní části okna.

Existuje také možnost, že máme v obou oknech zvýrazněná klíčová slova, která hledáme (tzv. „KWIC“ – Keyword In Context). Prvně se ujistíme, že máme spodní okno aktivní, poté vybereme možnost *Context Type* v nabídce *Display* (hlavní nabídka) a zvolíme možnost *Words*. Poté v menu *Sort* najedeme myší na možnost *Search Term* a klikneme na *1st Left*. V možnosti *Search Query* můžeme pro jeden francouzský termín

zadat několik českých ekvivalentů, které mají být vyhledány. Názornou ukázkou je obrázek č. 5.

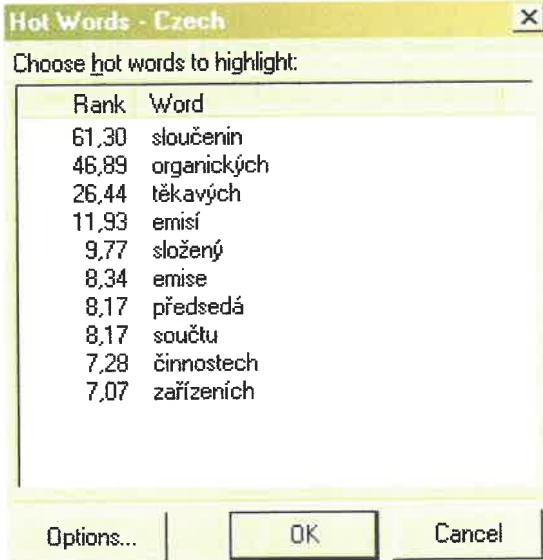


Obrázek č. 5: *Paralelní zobrazení klíčových slov (KWIC)*

Nesmíme se ale nechat zmást podvojným zvýrazněním klíčových slov, neboť není zaručeno, že si termín se svým překladem odpovídají v jednom konkordančním řádku.

Místo volby Search Query můžeme využít tabulku tzv. „Hot Words“, která nám sama nabídne nejčastěji se vyskytující termíny ve spodním okně (tedy v mém případě okno s českým překladem). Tuto tabulku získáme tak, že ve spodním okně (musí být aktivní) klikneme pravým tlačítkem myši a zvolíme možnost *Hot Words*. Pokud jsme předtím použili Search Query, je třeba tuto funkci vypnout a to tak, že zvolíme možnost

Clear Search Query. Nyní se nám nabízí tabulka s nejfrekventovanějšími slovy v českém, spodním okně, viz. obrázek č. 6.



Rank	Word
61,30	sloučenin
46,89	organických
26,44	těkavých
11,93	emisí
9,77	složený
8,34	emise
8,17	předseda
8,17	součtu
7,28	činnostech
7,07	zařízeních

Obrázek č. 6: *Tabulka „Hot Words“*

Máme možnost vybrat buď jen některá (s přidržením klávesy Ctrl) nebo všechna slova v dané tabulce a ty pak budou v konkordančních řádcích zvýrazněny. V menu *Options* lze nastavit počet vyhledaných „Hot Words“.

3.2.5 Frekvence slov

Program ParaConc nám nabízí dva základní druhy frekvenční statistiky. První je *korpusová* (tedy všech slov v ParaConcu nahraných) a druhá pak statistika *kolokační* (frekvence slov vázaných na vyhledané slovo), jak je uvedeno na <http://www.athel.com/para.html>.

Co se týče frekvence slov ve všech nahraných textech, máme dvě možnosti, jak tuto statistiku sledovat. Buď pomocí funkce *Frequency Order* (v menu *Frequency*), kdy

se nám zobrazí tabulka s termíny od nejfrekventovanějšího po nejméně frekventovaný termín. Lze vybrat jak frekvenci slov v jednom jazyku nebo v obou jazycích. Tato tabulka udává jak počet výskytů daných termínů, tak i procentuální zastoupení v textech. A nebo pak pomocí funkce *Alphabetical Order*, kdy se nám termíny řadí nikoliv podle frekvence výskytu, ale podle abecedy. Opět je možné vybrat jeden jazyk nebo oba a taktéž je zde uveden jak počet termínu tak procenta.

V druhém případě se pak jedná o frekvenci slov, která jsou „v okolí“ hledaného slova. Když nám program dané slovo vyhledá, pak stačí v menu *Frequency* zvolit *Collocate Frequency Data* a nabídne se nám tabulka, která obsahuje slova vyskytující se nalevo a napravo od hledaného slova. Kolik sloupců slov se nám ukáže závisí na tom, jakou možnost jsme zvolili ve *Frequency Options*. Pokud v poli *Collocate Span* udáme *1L - 1R*, pak se nám v tabulce frekvence objeví pouze jedno slovo nalevo a jedno napravo od hledaného termínu. Nanejvýš je možno zvolit *4L - 4R* (čtyři zleva – čtyři zprava). U každého slova je pak uveden počet výskytů, viz obr. 7.

Je třeba si uvědomit, že program nerozeznává singulár a plurál substantiv, tudíž pokud je u daného slova (které se nachází například vpravo od hledaného termínu) uveden počet 23, nemusí to být konečný výsledek, neboť se v textu může nacházet ten samý termín ještě v plurálu, ale třeba jen dvakrát a potom je uveden až na konci výčtu slov. V tomto případě není výsledný počet slov 23, ale 25. Tzn., že program při vyhodnocování takovéto statistiky pokládá singulár a plurál substantiva za dva různé termíny.

2-Left	1-Left	1-Right	2-Right
9 émissions	16 de	21 organiques	12 volatils
2 des	6 des	2 d'un	2 ayant
2 ou	2 comité	2 des	2 l'étiquetage
2 somme	2 différents	2 justifiant	2 par
2 un	2 tout		2 représentant
			2 représentants

Obrázek č. 7: Kolokační frekvenční statistika

Díky programu ParaConc lze zkoumat paralelní texty dvou či více jazyků způsobem, který ocení nejen autoři slovníků či jiných cizojazyčných příruček, ale i široká veřejnost, která má zájem analyzovat cizí jazyky či obohatit si své znalosti netradiční cestou.

3.3 Texty z oblasti „Ochrana zdraví“

Jak je již zmíněno v úvodu, pro analyzování francouzské a české terminologie v evropských legislativních textech jsem si vybrala oblast „Ochrana zdraví“. Jedná se o evropské legislativní akty jako směrnice, nařízení či rozhodnutí. Jejich seznam je uveden v tabulce č. 3., která obsahuje číslo dokumentu, pod kterým vystupuje v úředním věstníku Evropské unie na stránkách <http://europa.eu.int/eur-lex/>, dále pak rok vydání, číslo řady úředního věstníku a číslo úředního věstníku, podle kterých je možné text na výše uvedené adrese vyhledat. Nechybí ani počet normostran každého dokumentu.

Tabulka č. 3: Informace o analyzovaných textech

ČÍSLO DOKUMENTU	ROK VYDÁNÍ	ŘADA ÚŘ. VĚSTNÍKU	ČÍSLO ÚŘ. VĚSTNÍKU	POČ. NORMOSTRAN
31989L0618	1989	L	357	6
31992L0109	1992	L	370	7
31993R0302	1993	L	36	12
31993R1493	1993	L	148	4
31994D0385	1994	L	176	1
31995D0568	1995	L	324	1
31996D0158	1996	L	37	3
31996R1485	1996	L	188	2
31996D0469	1996	L	192	3
31997R0550	1997	L	85	9
31998D0291	1998	L	131	2
31998D2119	1998	L	268	9
31998L0079	1998	L	331	59
31999L0013	1999	L	85	27
32000R0141	2000	L	18	9
32000D0057	2000	L	21	5
32000R0847	2000	L	103	7
32000L0069	2000	L	313	12
32000L0070	2000	L	313	5
32001L0037	2001	L	194	16
32003L0012	2003	L	28	1
32003R1568	2003	L	224	12

V následující tabulce (tabulka č. 4) jsou uvedeny čísla a názvy analyzovaných textů jak v jazyce českém tak francouzském.

Tabulka č. 4: Názvy analyzovaných dokumentů

ČÍSLO DOKUMENTU	NÁZEV DOKUMENTU
31989L0618	Směrnice Rady o informování obyvatelstva o opatřeních na ochranu zdraví, která se mají použít, a o krocích, které je třeba učinit v případě radiální mimořádné situace. Directive du Conseil concernant l'information de la population sur les mesures de protection sanitaire applicables et sur le comportement à adopter en cas d'urgence radiologique.

<p>31992L0109</p>	<p>Směrnice Rady o výrobě určitých látek používaných k nedovolené výrobě omamných a psychotropních látek a o jejich uvádění na trh.</p> <p>Directive du Conseil relative à la fabrication et à la mise sur le marché de certaines substances utilisées pour la fabrication illicite de stupéfiants et de substances psychotropes.</p>
<p>31993R0302</p>	<p>Nařízení Rady o zřízení Evropského monitorovacího centra pro drogy a drogovou závislost.</p> <p>Règlement du Conseil portant création d'un observatoire européen des drogues et des toxicomanies.</p>
<p>31993R1493</p>	<p>Nařízení Rady o přepravě radioaktivních látek mezi členskými státy.</p> <p>Règlement du Conseil concernant les transferts de substances radioactives entre les États membres.</p>
<p>31994D0385</p>	<p>Rozhodnutí Komise o uvedení na trh produktu skládajícího se z geneticky modifikovaného organismu, osiva tabáku odrůdy ITB 1000 OX odolné vůči herbicidu.</p> <p>Décision de la Commission concernant la mise sur le marché d'un produit consistant en un organisme génétiquement modifié, à savoir des semences de la variété de tabac ITB 1000 OX, résistant aux herbicides.</p>
<p>31995D0568</p>	<p>Rozhodnutí Rady o uzavření dohody mezi Evropským společenstvím a Kolumbijskou republikou o prekurzorech a chemických látkách často používaných k nedovolené výrobě omamných nebo psychotropních látek.</p> <p>Décision du Conseil concernant la conclusion de l'accord entre la Communauté européenne et la république de Colombie relatif aux précurseurs et aux substances chimiques utilisés fréquemment pour la fabrication illicite de drogues ou de substances psychotropes.</p>
<p>31996D0158</p>	<p>Rozhodnutí Komise o uvedení na trh produktu složeného z geneticky modifikovaného organismu, osiva hybridní řepky tolerantní k herbicidu (Brassica napus L. oleifera Metzq. MS1Bn × RF1Bn).</p> <p>Décision de la Commission concernant la mise sur le marché d'un produit consistant en un organisme génétiquement modifié, à savoir des semences de colza hybride tolérant aux herbicides (Brassica napus L. oleifera Metzq. MS1Bn × RF1Bn).</p>
<p>31996R1485</p>	<p>Nařízení Komise, kterým se stanoví prováděcí pravidla ke směrnici Rady 92/109/EHS, pokud jde o prohlášení zákazníka o zvláštním použití určitých látek používaných k nedovolené výrobě omamných a psychotropních látek.</p> <p>Règlement de la Commission portant modalités d'application de la directive 92/109/CEE du Conseil en ce qui concerne les déclarations du client qui spécifient les usages de certaines substances utilisées dans la fabrication illicite de stupéfiants et de substances psychotropes.</p>

31996D0469	<p>Rozhodnutí Komise o zřízení Poradního výboru pro prevenci rakoviny. Décision de la Commission portant création d'un comité consultatif pour la prévention du cancer.</p>
31997R0550	<p>Nařízení Rady o akcích v souvislosti s HIV/AIDS v rozvojových zemích. Règlement du Conseil relatif aux actions dans le domaine du VIH/sida dans les pays en développement.</p>
31998D0291	<p>Rozhodnutí Komise o uvedení na trh geneticky modifikované jarní řepky olejnè (<i>Brassica napus</i> L. ssp. <i>oleifera</i>). Décision de la Commission concernant la mise sur le marché de semences de colza de printemps génétiquement modifié (<i>Brassica napus</i> L. ssp. <i>oleifera</i>).</p>
31998D2119	<p>Rozhodnutí Evropského parlamentu a Rady o zřízení sítě epidemiologického dozoru a kontroly přenosných nemocí ve Společenství. Décision du Parlement Européen et du Conseil instaurant un réseau de surveillance épidémiologique et de contrôle des maladies transmissibles dans la Communauté.</p>
31998L0079	<p>Směrnice Evropského parlamentu a Rady o diagnostických zdravotnických prostředcích in vitro. Directive du Parlement Européen et du Conseil relative aux dispositifs médicaux de diagnostic in vitro.</p>
31999L0013	<p>Směrnice Rady o omezování emisí těkavých organických sloučenin vznikajících při používání organických rozpouštědel při některých činnostech a v některých zařízeních. Directive du Conseil relative à la réduction des émissions de composés organiques volatils dues à l'utilisation de solvants organiques dans certaines activités et installations.</p>
32000R0141	<p>Nařízení Evropského parlamentu a Rady o léčivých přípravcích pro vzácná onemocnění. Règlement du Parlement Européen et du Conseil concernant les médicaments orphelins.</p>

32000D0057	<p>Rozhodnutí Komise o systému včasného varování a reakce pro účely prevence a kontroly přenosných nemocí.</p> <p>Décision de la Commission concernant le système d'alerte précoce et de réaction pour la prévention et le contrôle des maladies transmissibles.</p>
32000R0847	<p>Nařízení Komise, kterým se stanoví prováděcí pravidla ke kritériím pro stanovení léčivého přípravku jako léčivého přípravku pro vzácná onemocnění a definice pojmů „podobný léčivý přípravek“ a „klinická nadřazenost“.</p> <p>Règlement de la Commission établissant les dispositions d'application des critères de désignation d'un médicament en tant que médicament orphelin et définissant les concepts de «médicament similaire» et de «supériorité clinique».</p>
32000L0069	<p>Směrnice Evropského parlamentu a Rady o mezních hodnotách pro benzen a oxid uhelnatý v ovzduší.</p> <p>Directive du Parlement Européen et du Conseil concernant les valeurs limites pour le benzène et le monoxyde de carbone dans l'air ambiant.</p>
32000L0070	<p>Směrnice Evropského parlamentu a Rady, kterou se mění směrnice Rady 93/42/EHS s ohledem na zdravotnické prostředky obsahující stabilní deriváty z lidské krve nebo lidské plazmy.</p> <p>Directive du Parlement Européen et du Conseil modifiant la directive 93/42/CEE du Conseil en ce qui concerne les dispositifs médicaux incorporant des dérivés stables du sang ou du plasma humains.</p>
32001L0037	<p>Směrnice Evropského parlamentu a Rady o sbližování právních a správních předpisů členských států týkajících se výroby, obchodní úpravy a prodeje tabákových výrobků.</p> <p>Directive du Parlement Européen et du Conseil relative au rapprochement des dispositions législatives, réglementaires et administratives des États membres en matière de fabrication, de présentation et de vente des produits du tabac.</p>
32003L0012	<p>Směrnice Komise o změně klasifikace prsních implantátů v rámci směrnice 93/42/EHS o zdravotnických prostředcích.</p> <p>Directive de la Commission concernant la reclassification des implants mammaires dans le cadre de la directive 93/42/CEE relative aux dispositifs médicaux.</p>
32003R1568	<p>Nařízení Evropského Parlamentu a Rady o podpoře boje proti chorobám spojeným s chudobou (HIV/AIDS, tuberkulóza a malárii) v rozvojových zemích.</p> <p>Règlement du Parlement Européen et du Conseil relatif à l'aide en faveur de la lutte contre les maladies dues à la pauvreté (VIH/sida, tuberculose et paludisme) dans les pays en développement.</p>

3.4 Klíčová slova z oblasti „Ochrana zdraví“

Podle údajů normostran tedy vyplývá, že k analýze klíčových slov z oblasti „Ochrana zdraví“ jsem měla k dispozici celkem 212 stran textu, kdy celkový počet slov všech textů je pro český jazyk 55 317 a pro jazyk francouzský 75 267.

Tyto dokumenty v elektronické podobě jsem analyzovala pomocí programu ParaConc a to způsobem, jež je uveden v kapitole 3.2. Nejprve bylo ale zapotřebí si texty pozorně přečíst a vybrat z nich klíčová slova pro danou oblast. Tato slova jsou samozřejmě součástí mnoha kolokací (viz kap. 3.2.3), které se staly taktéž předmětem mé analýzy.

Pokud jde o slovní druhy, tak vybraná klíčová slova jsou především substantiva, pouze v jednom případě se jedná o adjektivum. Je třeba si uvědomit, že klíčové slovo není vždy hlavou celé kolokace, může být pouze její součástí. Tzv. „hlava“ je základní prvek celé kolokace, který určuje její gramatickou povahu.

Promítneme-li vybrané termíny v číslech, pak celkový počet analyzovaných termínů (jak klíčových slov tak i jejich kolokací) činí 158. Samotných klíčových slov je pak 26 (tabulka č. 5).

Cílem této kapitoly je přiblížit vybranou francouzskou terminologii jak z hlediska morfologického, tak i z hlediska jejich překladů do českého jazyka. Klíčové termíny jsem rozdělila do čtyř kategorií podle homogenity jejich překladů a to na skupinu slov s jednoznačným překladem, skupinu slov překládaných několika synonymy, pak skupinu slov přeložených nepřesně, skupinu slov s chybným překladem a skupinu slov, u kterých jsem pozorovala nesrovnalosti ve francouzských textech nikoliv u jejich překladů. U každého analyzovaného klíčového slova je uvedena tabulka jeho kolokací a počet výskytů těchto kolokací v textech.

Tabulka č. 5: Seznam vybraných klíčových slov

Klíčové slovo	Slovní druh
affection	s.f.
cancer	s.m.
colza	s.m.
comité	s.m.
composé	s.m.
dispositif	s.m.
dissémination	s.f.
drogue	s.f.
énergie	s.f.
épidémie	s.f.
gène	s.m.
herbicide	s.m.
maladie	s.f.
médicament	s.m.
monoxyde de carbone	s.m.
nucléaire	adj.
organisme	s.m.
organisme 2	s.m.
pollution	s.f.
protection	s.f.
santé	s.f.
solvant	s.m.
substance	s.f.
tabac	s.m.
urgence	s.f.
VIH	s.m.

3.4.1 Klíčová slova a jejich kolokace přeložené chybně

V této skupině bych uvedla klíčová slova „maladie“ a „colza“, jelikož jsem při jejich analyzování narazila na chybné překlady a to u kolokace „maladie cardiaque“ a „maladie cardio-vasculaire“ a při překladu samotného klíčového slova „colza“.

Termín „maladie“ je substantivum rodu ženského, jehož celkový počet výskytů v textech je 119. Z toho samotný termín bez kolokací se zde vyskytuje jen 42krát. Je

překládán především jako „nemoc“, ale v pěti případech jako „onemocnění“. Podle Velkého francouzsko-českého slovníku (1992 : 54) je překlad tohoto termínu následující: nemoc, choroba, onemocnění. Podle francouzského výkladového slovníku Larousse (2004 : 489) definice termínu zní: trouble, dérangement de la santé physique, du comportement, etc.

Pokud jde o kolokace tohoto termínu uvedené v tabulce č. 6, pak se tady objevují určité nesrovnalosti, přičemž jeden překlad je chybný.

Tabulka č. 6: Kolokace termínu „maladie“

Kolokace	Česky - 1	Česky - 2
maladie (42) ²	nemoc	onemocnění
maladie transmissible (53) ³	přenosná nemoc	
maladie rare (13)	vzácná nemoc	
maladie due à la pauvreté (5)	nemoc spojená s chudobou	
maladie cardiaque (1)	onemocnění srdce	
maladie cardio-vasculaire (1)	srdeční onemocnění	
maladie héréditaire (1)	dědičné onemocnění	
maladie d'origine alimentaire (1)	nemoc přenášená potravou	
maladie d'origine hydrique et environnementale (1)	nemoc přenášená vodou a nemoc závislá na prostředí	
maladie transmissible par des agents non conventionnels (1)	nemoc přenášená nekonvenčními činiteli	

Chybná je kolokace „maladie cardio-vasculaire“, jež je přeložena skoro stejně jako kolokace „maladie cardiaque“ - „onemocnění srdce“ a to jako „srdeční

² U klíčových slov v tabulkách je v závorce uveden počet výskytů samotného termínu bez jeho kolokací.

³ U kolokací v tabulkách je v závorce uveden počet výskytů celých těchto kolokací.

onemocnění“. Chybný překlad dokazuje Velký francouzsko-český slovník, kde je adjektivum „cardiaque“ přeloženo jako „srdeční“ nebo „kardiální“, zatímco „cardio-vasculaire“ odpovídá českému ekvivalentu „kardiovaskulární“. Z lékařského hlediska je pak rozdíl mezi těmito dvěma adjektivy podle Nového akademického slovníku cizích slov (2005 : 392) následující: „kardiální“ – srdeční, kdežto „kardiovaskulární“ – týkající se srdce a cév. Je tedy jisté, že překladatelé tento rozdíl opomenuli a nabídli nám zcela totožný překlad těchto dvou rozdílných termínů.

Zajímavá je pak i kolokace „maladie due à la pauvreté“, kdy je překládána jako „nemoc spojená s chudobou“. Přičemž termín „dû à qc“ znamená „způsobený, zaviněný něčím“ (Velký francouzsko-český slovník, 1992 : 506). Pokud bychom chtěli nalézt přesný francouzský ekvivalent pro spojení „spojená s..“, pak by se jednalo o „liée à qc“. Toto spojení se v textech vyskytuje pouze jednou.

Jako poslední zajímavý překlad v případě slova „maladie“ bych chtěla uvést poslední tři kolokace tohoto slova uvedené v tabulce č. 6.

Zprv se jedná o český ekvivalent pro „maladie d'origine alimentaire, hydrique et environnementale“ a „maladie transmissible par des agents non conventionnels“, kdy se v českých textech objevuje stejný překlad a to „nemoc přenášená potravou, vodou a nemoc závislá na prostředí“ a „nemoc přenášená nekonvenčními činiteli“. Dalo by se uvažovat o případném překladu kolokace „maladie d'origine alimentaire“, který by mohl znít: „nemoc potravinového původu“, ale asi by to pak neznělo příliš česky a už vůbec ne v případě, kdybychom takový překlad použili pro „maladie d'origine hydrique et environnementale“ (bylo by pak možné to přeložit třeba jako „nemoc vodního původu“ apod., ale to je pro český jazyk nepřijatelné). Z toho vyplývá, že se překladatelé takovému „konfliktu“ vyhnuli a použili překlad „nemoc přenášená...(čím)“.

A zadruhé je zajímavé, že „maladie d'origine environnementale“ je přeloženo jako „nemoc závislá na prostředí“ nikoliv jako „nemoc přenášená prostředím“. Uvedený překlad má své opodstatnění, neboť překladatelé nejspíš vzali v úvahu to, že samo životní prostředí by v případě, kdyby bylo v pořádku, žádnou nemoc nepřenášelo. Podle mého názoru by bylo možné použít i „nemoc způsobená prostředím“.

Druhým velmi zajímavým termínem z hlediska jeho českého překladu je termín „colza“, jehož celkový počet v textech činí 13, z toho 3 výskyty náleží jen samotnému termínu. Tento termín je substantivum rodu mužského a v textech je překládán většinou jako „řepka olejka“, někdy jako „řepka olejná“ nebo jen „řepka“. Na základě mého výzkumu jsem zjistila, že „colza“ není „řepka“, ale „řepice“ a navíc, řepka (*Brassica napus* L.) a řepice (*Brassica rapa* L.) jsou dvě různé rostliny, i když velmi podobné. (http://genbank.vurv.cz/genetic/resources/documents/documents_cz.htm). Navíc jsem na webových stránkách <http://isap.vlada.cz/> zjistila, že „colza“ je závazný termín pro překlad a v češtině zní „řepice“, zatímco „řepka“ má francouzský ekvivalent „navette“. Z toho tedy vyplývá, že překladatelé tuto důležitou informaci opomenuli. V příloze pak uvádím Směrnici Komise ze dne 27. června 1979, kterou se mění botanické názvy některých rostlin včetně řepky a řepice vzhledem k vývoji vědeckých a technických poznatků.

Pokud jde o kolokace tohoto slova (tabulka č. 7), jsou překládány chybně, jelikož chybný je překlad klíčového slova, jež je ve všech kolokacích obsažen.

Tabulka č. 7: *Kolokace termínu „colza“*

Kolokace	Česky - 1	Česky - 2
colza (3)	řepka olejka	řepka
colza hybride (2)	hybridní řepka	
colza non génétiquement modifiée (2)	geneticky nemodifikovaná řepka	
colza génétiquement modifiée (2)	geneticky modifikovaná řepka	
colza de printemps (2)	jarní řepka olejná	
semences de colza (2)	osiva řepky olejně	

3.4.2 Klíčová slova a jejich kolokace s nepřesným překladem

Do této skupiny bych zařadila termín „**pollution**“, což je substantivum rodu ženského, v textech se vyskytuje celkem 10krát, z toho 6krát jen samotný termín bez jeho kolokací. Mezi kolokace tohoto klíčového slova jsem zařadila i termín „polluant“, což sice není jeho kolokace, ale jedná se o slovo se stejným morfologickým základem, tudíž jsem tyto dva termíny uvedla dohromady (tabulka č. 8). Termín „polluant“ se pak v textech vyskytuje celkem 12krát, z toho 9krát bez kolokací a to jak ve formě substantiva rodu mužského tak i ve formě adjektiva.

Na překladech těchto dvou termínů je zajímavé to, že termín „pollution“ je překládán nejčastěji jako „znečištění“, ale objevuje se i překlad „znečišťující látka“. Podle Velkého francouzsko-českého slovníku (1992 : 306) je překlad tohoto slova jen „znečištění“. Další dva překlady totiž odpovídají podle textů francouzskému ekvivalentu „polluant“. Přesto je takto přeložen i termín „pollution“. Pro představu uvádím širší kontext, kde je tento překlad použit.

- „.....des schémas probables de répartition de la **pollution de l'air ambiant** et de l'exposition potentielle de la population.“

„.....pravděpodobný rozptyl **látek znečišťujících ovzduší** a možné vystavení obyvatelstva zvýšeným koncentracím.

- „.....correspondre aux plus hauts niveaux de **pollution** avec les résultats fournis par les inventaires des émissions...“

„.....typická pro nejvyšší hodnoty **znečišťujících látek**, spojit s informací z emisních inventur....“

V prvním případě nás může upoutat již slovo „répartition“, jež je překládáno jako „rozptyl“, ačkoliv podle Velkého francouzsko-českého slovníku (1992 : 439) je tento termín přeložen jako „rozdělení, rozložení, rozvržení“. Kdybychom chtěli tento úsek věty přeložit podle „přesných“ překladů ve slovníku, jedna z variant by mohla být např. „pravděpodobná schémata vyznačující jak je znečištění ovzduší rozloženo“.

Překladaelé se ale nejspíš přiklonili k volnějšimu překladu, který sice není příliš přesný, ale vystihuje podstatu věci a zní více „česky“.

V druhém případě by pak místo překladu „znečišťujících látek“ mohl být překlad „znečištění“, zněl by přirozeně a byl by přesnější.

Tabulka č. 8 : Kolokace termínu „pollution“

Kolokace	Česky - 1	Česky - 2
pollution (6)	znečištění	znečišťující látka
pollution atmosphérique (2)	znečišťování ovzduší	
pollution de l'atmosphère (1)	znečištění ovzduší	
échantillonneur mesurant la pollution (1)	měřicí stanice	
polluant (9)	znečišťující látka	
polluant – adj. (1)	znečišťující	
polluant atmosphérique (2)	látka znečišťující ovzduší	

3.4.3 Klíčová slova a jejich kolokace s nesrovnalostmi ve francouzských textech

Zde se jedná o dva termíny a to „VIH“ a „comité“, kdy jejich překlady odpovídají všem slovníkům, které jsem měla při analýze k dispozici a odpovídají i databázi ISAP, avšak nesrovnalosti se objevují v originálních francouzských textech.

Klíčové slovo „VIH“ je rodu mužského a jedná se o tzv. „sigle“, což znamená zkratka z počátečních písmen. Tato zkratka znamená HIV, což je anglická zkratka pro

„Human Immunodeficiency Virus (Nový akademického slovník cizích slov (2005: 309), kterou čeština převzala bez jakýchkoliv změn. Ve francouzštině byla tato zkratka změněna. Zůstává pořád zkratkou z počátečních písmen, avšak ve francouzštině jsou delší adjektiva či přívlastky postaveny až za substantivum, tudíž VIH je zkratkou pro název „Virus d'immunodéficience humaine“. Stejná změna nastává i u zkratky „Aids“ (Acquired Immune Deficiency Syndrome), která se ve francouzštině užívá ve formě „Sida“ (Syndrome immunodéficitaire acquis) a v textech je často se zkratkou VIH pojena (viz tabulka č. 9).

Jak bylo již řečeno v tomto případě jsem u překladů nepozorovala žádnou nesrovnalost, je jen zajímavé, že se ve francouzských textech vyskytla pětkrát zkratka HIV, i když v ostatních 35 případech je v textech uvedena zkratka VIH.

Tabulka č. 9: Kolokace termínu „VIH“

Kolokace	Česky - 1
VIH(13)	HIV
VIH/sida (19)	HIV/AIDS
transmission du VIH/sida (3)	přenos HIV/AIDS

Pokud jde o klíčové slovo „comité“, je to substantivum rodu mužského a v textech se vyskytuje celkem 163krát, z toho bez kolokací 119krát. Je překládán jednoznačně jako „výbor“. V případě tohoto termínu jsem nenalezla žádnou nesrovnalost v jeho překladech do českého jazyka, jen se v textech objevuje nejednotnost u velkých a malých počátečních písmen názvů některých výborů. Jedná se o „comité scientifique et technique“ – „Výbor pro vědu a techniku“ a „comité des médicaments orphelins“ – „Výbor pro léčivé přípravky pro vzácná onemocnění“, kdy jsou ve francouzských názvech výborů malá počáteční písmena, i přesto, že v databázi

ISAP jsou uvedeny s velkými počátečními písmeny. Všechny kolokace tohoto klíčového slova jsou uvedeny v tabulce č. 10.

Tabulka č. 10: Kolokace termínu „comité“

Kolokace	Česky - 1
comité (119)	výbor
comité scientifique et technique(2)	Výbor pro vědu a techniku
Comité économique et social (12)	Hospodářský a sociální výbor
comité consultatif (3)	poradní výbor
comité des médicaments orphelins (6)	Výbor pro léčivé přípravky pro vzácná onemocnění
Comité europ. de lutte antidrogues (1)	Evropský výbor pro boj s drogami
Comité europ. de normalisation (2)	Evropský výbor pro normalizaci
Comité des régions (3)	Výbor regionů
Comité "Dispositifs médicaux" (2)	Výbor pro zdravotnické prostředky
Comité "Normes et règles techniques" (1)	Výbor pro normy a technické předpisy
comité scientifique (12)	vědecký výbor

3.4.4 Klíčová slova a jejich kolokace překládané více synonymy

Do této skupiny jsem zařadila slova, jež jsou překládány více českými ekvivalenty než jedním. Jedná se o termíny:

affection

composé

dissémination

drogue

herbicide

médicament

monoxyde de carbon

organisme 2

santé

tabac

urgence.

Podíváme-li se na první termín výčtu „**affection**“, je to substantivum rodu ženského, jež se v textech vyskytuje méně často než jeho synonymum „maladie“ a to 33krát, z toho pak samotný termín 28krát. V textech je překládán jako „onemocnění“ pouze v jednom případě pak jako „nemoc“. Velký francouzský slovník (1992 : 82) uvádí překlad: afekce, chorobný stav, choroba; výkladový slovník Larousse (2004 : 15) uvádí definici: état maladif. V obou slovnících je uvedena zkratka „med.“, což tedy znamená, že termín „affection“ se používá především v lékařství.

Podíváme-li se na kolokace termínu „affection“, vypořadujeme zde to, že pro „affection mettant la vie en danger“ a „affection entraînant une menace pour la vie“ je v českých textech uveden jen jeden překlad a to „život ohrožující onemocnění“. Pro druhou uvedenou kolokaci bychom mohli použít například překlad „onemocnění

způsobující ohrožení života“, tato možnost je ale delší a i přes to, že vyjadřuje podstatu věci, přikláním se i já k původní a kratší formě překladu.

Tabulka č. 11: Kolokace termínu „affection“

Kolokace	Česky - 1	Česky - 2
affection (28)	onemocnění	nemoc
affection mettant la vie en danger, très invalidante ou grave et chronique (2)	život ohrožující, závažně invalidizující nebo závažná a chronická onemocnění	
affection entraînant une menace pour la vie ou une invalidité chronique (2)	život ohrožující nebo chronicky invalidizující onemocnění	
affection rare (1)	vzácné onemocnění	

Dále následuje klíčové slovo „**composé**“, což je substantivum rodu mužského a v textech se vyskytuje celkem 36krát, z toho 10krát bez studovaných kolokací. Je překládáno jako „sloučenina“ nebo „látka“, což odpovídá všem slovníkům, se kterými jsem pracovala. V textech vystupuje taktéž jako adjektivum a v tomto případě je překládán jako „složený“. Pro přehled uvádím tabulku č. 12, kde lze vyzorovat, kdy jsou přeloženy jednotlivé kolokace tohoto termínu českým ekvivalentem „sloučenina“ a kdy jako „látka“. U kolokace „composé organique volatil“ je ve francouzských textech uvedena zkratka „COV“, která se pak sama v textech nachází 13krát. V českých textech se ekvivalent této zkratky neuvádí a je překládána celým názvem.

Tabulka č. 12: *Kolokace termínu „composé“*

Kolokace	Česky - 1	Česky - 2
composé (10)	sloučenina	látka
composé organique volatil (12) – COV (13)	těkavá organická sloučenina	
Composé(3)	složený	
composé biologique (1)	biologická látka	
composé organique (10)	organická sloučenina	

Klíčové slovo „**dissémination**“ je substantivum rodu ženského a v textech se vyskytuje celkem 6krát, 3krát pak samotný termín bez kolokací. Je přeloženo dvěma způsoby a to buď jako „uvolňování“ anebo jako „rozšiřování“. Dá se konstatovat, že tyto překlady jsou užity půl na půl, neboť když je termín „dissémination“ v kolokaci „dissémination volontaire“, pak je překládán jako „uvolňování“, což je z celkového počtu výskytu 3krát. V ostatním třech případech je pak tento termín přeložen jako „rozšiřování“, viz tabulka č. 13.

Tabulka č. 13: *Kolokace termínu „dissémination“*

Kolokace	Česky - 1	Česky - 2
dissémination (3)	uvolňování	rozšiřování
dissémination volontaire (3)	záměrné uvolňování	

Pokud jde o termín „**drogue**“, je to substantivum rodu ženského, vyskytuje se v textech celkem 38krát s tím, že samotný termín bez studovaných kolokací pouze 5krát. V textech je překládán jako „droga“ či „omamná látka“. Zajímavější než samotný termín jsou jeho kolokace, viz tabulka č. 14.

Tabulka č. 14: Kolokace termínu „drogue“

Kolokace	Česky - 1	Česky - 2	Česky - 3
drogue (5)	droga	omamná látka	
drogues et des toxicomanies (11)	drogy a drogová závislost	zneužívání drog	toxikomanie
phénomène de la drogue (7)	problematika drog	jev drog	oblast drog
fabrication illicite de drogues (4)	nedovolená výroba drog		
lutte contre la drogue (4)	boj s drogami		
Observatoire européen des dr. et des toxicomanies(4)	Evropské monitorovací centrum pro drogy		
demande de drogue(s) (3)	poptávka po drogách		

První kolokace „drogues et des toxicomanies“ je překládána třemi způsoby a to „drogy a drogová závislost“ (což je překlad nejpřesnější), dále pak jako „zneužívání drog“ a „toxikomanie“, což jsou překlady volnější, ale pořád vystihují smysl překládaného termínu.

Kolokace „phénomène de la drogue“ je překládán taktéž třemi způsoby, kdy nejpřesnější je „jev drog“ (podle Velkého francouzsko-českého slovníku (1992 : 271) totiž termín „phénomène“ znamená „jev, fenomén“), ale ten se v textu vyskytuje jen jednou. Další dva překlady, které se oba v textech vyskytují třikrát zní: „problematika drog“ a „oblast drog“. Tyto překlady jsou v textech použity častěji v závislosti na kontextu, neboť překlad „jev drog“ by v určitých situacích nezněl příliš „česky“.

Dalším pozorováním je nikoliv překlad, ale samotné nesrovnalosti v originálních francouzských textech, kdy se instituce „Observatoire européen des drogues et des toxicomanies“ vyskytuje jak s malým tak i velkým počátečním písmenem. Podle webové stránky <http://isap.vlada.cz>, jež obsahuje databázi závazných termínů pro

překlady právních předpisů ES, je tato instituce s velkým počátečním písmenem. Z toho tedy vyplývá, že i v originálních textech se mohou objevit chyby.

V neposlední řadě bych pak chtěla uvést, že kolokace „demande de drogue(s)“ je vždy překládána jako „poptávka po drogách“ i přesto, že ve francouzštině lze použít jak singulár tak plurál termínu „drogue“. V češtině se výraz „poptávka po droze“ nepoužívá.

Dalším termínem je „**herbicide**“, což je taktéž substantivum rodu mužského a v textech se vyskytuje celkem 14krát, jednou pak bez kolokací. Je překládán jako „herbicid“ nebo „prostředek proti plevelu“, což odpovídá jak překladu závazných termínů na <http://isap.vlada.cz/>, kde je tento termín překládán jako „herbicid“, tak i Velkému francouzsko-českému slovníku (1992 : 741), kde se objevuje překlad „prostředek ničící n. hubící plevel“. Překlady kolokací (tabulka č. 15) jsou také bez nejasností.

Tabulka č. 15: Kolokace termínu: „herbicide“

Kolokace	Česky - 1	Česky - 2
herbicide (1)	herbicid	prostředek proti plevelu
herbicide chimique (3)	chemický herbicid	
résistant à herbicide (6)	odolný vůči herbicidu	
tolérant aux herbicides (4)	tolerantní k herbicidům	

Termín „**médicament**“ je překládán také bez jakýchkoliv nejasností, ale třemi různými synonymy. Jedná se o substantivum rodu mužského a v textech se vyskytuje celkem 152krát (77krát bez kolokací). Je překládán jako „léčivý přípravek“, „lék“ nebo „léčivo“, v závislosti na tom, v jaké kolokaci se vyskytuje, viz tabulka č. 16.

Tabulka č. 16: Kolokace termínu „médicament“

Kolokace	Česky - 1	Česky - 2	Česky - 3
médicament (77)	léčivý přípravek	lék	léčivo
médicament orphelin (55)	léčivý přípravek pro vzácná onemocnění		
médicament à usage humain (3)	humánní léčivý přípravek		
médicament similaire (8)	podobný léčivý přípravek		
médicament orphelin autorisé (2)	registrovaný léčivý přípravek		
médicament dérivé du sang ou du plasma humains (2)	léčivý přípravek pocházející z lidské krve nebo lidské plazmy		
Agence européenne pour l'évaluation des médicaments (EMA) (3)	Evropská agentura pro hodnocení léčivých přípravků (EMA)		
accès de tous aux médicaments (2)	všeobecná dostupnost léků		

Klíčové slovo „**monoxyde de carbone**“, je substantivum rodu mužského a počet jeho výskytů v textech činí 47 (23 jen samotný termín bez kolokací). Je překládán jednoznačně jako oxid uhelnatý, zajímavější je pak jeho kolokace (tabulka č. 17) „teneur en monoxyde de carbone des cigarettes“, jež je překládána jako „obsah oxidu uhelnatého v cigaretách“ nebo taky „úroveň“ nebo „míra“ ox. uhelnatého v cigaretách. Ve slovnících se objevuje jen překlad „obsah“, avšak dalo by se říci, že všechna tři

slova (obsah, míra, úroveň) jsou vlastně v tomto případě synonyma, takže překladatelé jen využili rozmanitosti českého jazyka.

Tabulka č. 17: Kolokace termínu „monoxyde de carbone“

Kolokace	Česky - 1	Česky - 2	Česky - 3
monoxyde de carbone (23)	oxid uhelnatý		
teneur en monoxyde de carbone des cigarettes (11)	úroveň oxidu uhelnatého v cigaretách	obsah oxidu uh. v cigaretách	míra oxidu uh. v cigaretách
benzène et le monoxyde de car. dans l'air ambiant (13)	benzen a ox. uhelnatý v ovzduší		

Pokud jde o klíčové slovo „**organisme 2**“, je oddělen číslem 2 od termínu „organisme“, jelikož tyto dva termíny mají zcela odlišný význam. Termín „organisme 2“ je překládán několika synonymy jako „orgán“, „subjekt“ nebo „organizace“. Jedná se o substantivum rodu mužského a v textech se vyskytuje celkem 149krát, z toho bez kolokací jen 32krát. Kolokace jsou uvedeny v tabulce č. 18. Všechny překlady jsou v souladu se slovníky, které jsem k analýze použila.

Tabulka č. 18: Kolokace termínu „organisme 2“

Kolokace	Česky - 1	Česky - 2	Česky - 3
organisme 2 (32)	orgán	subjekt	organizace
organisme de radiodiffusion (1)	rozhlas		
organisme notifié (104)	oznámený subjekt		
organisme compétent (5)	příslušný subjekt		
organisme de protection de l'environnement (1)	organizace na ochranu životního prostředí		
organisme de droit privé (1)	soukromoprávní subjekt		
organisme international (4)	mezinárodní orgán	mezinárodní subjekt	
organismes gouvernementaux ou non gouvernementaux (1)	vládní a nevládní subjekty		

Klíčové slovo „santé“ je substantivum rodu ženského a v textech se objevuje celkem 163krát, z toho pak samotný termín bez kolokací jen 64krát. Je překládáno dvěma různými ekvivalenty a to „zdraví“ nebo „zdravotní stav“, v některých kolokacích je pak překládáno i jako adjektivum „zdravotní“. Všechny překlady odpovídají slovníkům, se kterými jsem pracovala a nevyskytuje se v nich žádná nesrovnalost. Jak jsou kolokace tohoto klíčového slova překládány, je znázorněno v tabulce č. 19. Termín „santé“ je také součástí kolokace „protection de la santé“ vyskytující se v textech 25krát, tu jsem ale zařadila pod klíčový termín „protection“, který je uveden v kapitole 3.4.5.

Tabulka č. 19: Kolokace termínu „santé“

Kolokace	Česky - 1	Česky - 2
santé (64)	zdraví	zdravotní stav
santé publique (58)	veřejné zdraví	lidské zdraví
santé humaine (24)	lidské zdraví	
ministre de la santé (4)	ministr zdravotnictví	
santé d'un patient (3)	zdravotní stav pacienta	
santé des personnes (2)	zdraví jednotlivců	lidské zdraví
santé de la population (2)	zdraví obyvatelstva	
santé des consommateurs (1)	zdraví spotřebitele	
santé sexuelle et génésique (2)	sexuální a reprodukční zdraví	
santé sexuelle et reproductrice (1)	sexuální a reprodukční zdraví	
porter atteinte à la santé (2)	způsobit zdravotní újmu	

Dalším termínem z výčtu je „**tabac**“, jež je mužského rodu a v textech se vyskytuje celkem 81krát (14krát pak samotný termín).

Kolokace termínu „tabac“ (tabulka č. 20) mají většinou jednoznačný překlad, výjimku pak tvoří tři kolokace. „Tabac à usage orale“, což je překládáno jako „tabák pro orální užití“ nebo „tabákový výrobek, který se nekouří“, dále „tabagisme actif ou passif“ je překládán jako „aktivní a pasivní kouření“ nebo „aktivní a pasivní užívání cigaret“. U této kolokace je ale zajímavé, že slovo „tabagisme“ znamená podle Velkého francouzsko-českého slovníku (1992 : 588) „otrava tabákem nebo nikotinem“, podle výkladového slovníku Larousse (2004 : 792) zní definice tohoto termínu „intoxication due à l'abus du tabac“. Podle toho by se tedy dalo říci, že překlad není až tak bezchybný. Avšak ve Francouzsko-českém všeobecném slovníku (1997 – 2000, tabagisme) je uvedena kolokace „lutte contre le tabagisme“ a je přeložena jako „boj proti kouření“. Z toho tedy lze vyvodit, že se termín „tabagisme“ překládá jako „kouření“. V poslední řadě bych pak chtěla uvést kolokaci „plante de tabac“, která je

překládána jako „tabáková rostlina“ nebo „odrůda tabáku“, s tím, že druhý překlad neodpovídá slovníkům, které mám k dispozici, jelikož odrůda je i podle databáze ISAP překládána jako „variété“.

Tabulka č. 20: *Kolokace termínu „tabac“*

Kolokace	Česky - 1	Česky - 2
tabac (14)	tabák	
tabac à usage orale (6)	tabák pro orální užití	tabákový výrobek, který se nekouří
tabac à rouler (2)	tabák na balení cigaret	
produits du tabac (52)	tabákové výrobky	
tabagisme actif ou passif (2)	aktivní nebo pasivní kouření	aktivní nebo pasivní užívání cigaret
conséquences du tabagisme sur la santé (2)	zdravotní důsledky kouření	
tabac brut (1)	surový tabák	
plante de tabac (2)	odrůda tabáku	tabáková rostlina

Poslední klíčové slovo výčtu je substantivum rodu ženského „**urgence**“, jež je přeloženo v textech jako „opatření,“ nebo „mimořádná situace“ a v textech se vyskytuje celkem 52krát, bez kolokací 12krát. Všechny překlady jsou bez nejasností a souhlasí i s databází ISAP. Kolokace „*measure d'urgence*“ je překládána jako „havarijní opatření“ anebo „ochranné opatření“, kdy se druhý překlad v textech užívá i pro kolokaci „*measure de protection*“ viz str. 58. Všechny kolokace a jejich překlady lze pozorovat v tabulce č. 21.

Tabulka č. 21: Kolokace termínu „urgence“

Kolokace	Česky - 1	Česky - 2
urgence (12)	opatření	mimořádná situace
urgence radiologique (28)	radiační mimořádná situace	
en fonction de l'urg. de la question en cause (9)	s ohledem na naléhavost věci	podle naléhavosti věci
mesure d'urgence (2)	havarijní opatření	ochranné opatření
mesure d'urgence prévues pour alerter (1)	opatření k varování	

3.4.5 Klíčová slova a jejich kolokace přeložené jednoznačně

V této skupině jsou uvedeny klíčová slova a jejich kolokace, které jsou přeloženy bezchybně a jednoznačně. Jejich překlad odpovídá jak slovníkům, s nimiž jsem pracovala, tak i databázi ISAP. Není tedy třeba je podrobněji rozebírat a jsou uvedeny pouze v přehledné tabulce spolu s počtem výskytů v závorce a jejich českými ekvivalenty.

Jedná se o tato klíčová slova:

cancer, s.m., tabulka č. 22

dispositif, s.m, tabulka č. 23

énergie, s.f., tabulka č. 24

épidémie, s.f., tabulka č. 25

gène, s.m., tabulka č. 26

nucléaire, adj., tabulka č. 27

organisme, s.m., tabulka č. 28

protection, s.f., tabulka č. 29

solvant, s.m., tabulka č. 30

substance, s.f., tabulka č. 31

Tabulka č. 22: Kolokace termínu „cancer“

Kolokace	Česky - 1
cancer (9)	rakovina
prévention du cancer (7)	prevence rakoviny
l'Europe contre le cancer (3)	Evropa proti rakovině
programme d'action contre le cancer (1)	akční plán boje proti rakovině
plan d'action contre le cancer (2)	akční plán boje proti rakovině

Tabulka č. 23: Kolokace termínu „dispositif“

Kolokace	Česky - 1
dispositif (341)	prostředek
dispositif médical (35)	zdravotnický prostředek
dispositif médical de diagnostic (24)	diagnostický zdravotnický prostředek
dispositif destiné à l'évaluation des performances (9)	prostředek pro hodnocení funkční způsobilosti
dispositifs destinés à des autodiagnostic (13)	prostředky pro sebetestování
classification des implants mammaires comme dispositifs médicaux de la classe III (2)	klasifikace prsních implantátů jako zdravotnické prostředky třídy III

Tabulka č. 24: Kolokace termínu „énergie“

Kolokace	Česky - 1
énergie (2)	energie
énergie électrique (1)	elektrická energie
Communauté europ. de l'énergie atomique (2)	Evropské společenství pro atomovou energii
Agence internationale de l'énergie atomique (AIEA) (1)	Mezinárodní agentura pro atomovou energii (MAAE)
énergie gazeuze (1)	plyn
énergie hydraulique (1)	hydraulická energie
énergie pneumatique (1)	pneumatická energie

Tabulka č. 25: *Kolokace termínu "épidémie"*

Kolokace	Česky - 1
épidémie (7)	epidemie
épidémie du VIH/sida (4)	epidemie HIV/AIDS
gravité de l'épidémie (2)	závažnost epidemie
impact de l'épidémie sur... (2)	dopad epidemie na....
expansion de l'épidémie (3)	rozšiřování epidemie

Tabulka č. 26: *Kolokace termínu „gène“*

Kolokace	Česky - 1
gène (6)	gen
transfert du gène de résistance à... (2)	přenos genu tolerance k...
gène barnas issu (1)	gen pro barnasu
gène bar issu (2)	bar-gen
gène neo issu (2)	gen neo
gène barstar issu (1)	gen barstar

Tabulka č. 27: *Kolokace termínu „nucléaire“*

Kolokace	Česky - 1
nucléaire (1)	jaderný
accident nucléaire (2)	jaderná havárie
installation nucléaire (1)	jaderné zařízení
réacteur nucléaire (1)	jaderný reaktor
combustible nucléaire (2)	jaderné palivo
matière nucléaire (2)	jaderný materiál

Tabulka č. 28: Kolokace termínu „organisme“

Kolokace	Česky - 1
organisme (1)	organismus
micro-organisme (1)	mikroorganismus
organisme génétiquement modifié (7)	geneticky modifikovaný organismus

Tabulka č. 29: Kolokace termínu „protection“

Kolokace	Česky - 1
protection (18)	ochrana
protection sanitaire (10)	ochrana zdraví
protection de la santé (25)	ochrana zdraví
protection de l'environnement (3)	ochrana životního prostředí
protection contre... (4)	ochrana před...
mesure de protection (12)	ochranné opatření
radioprotection (6)	radiační ochrana

Tabulka č. 30: Kolokace termínu „solvant“

Kolokace	Česky - 1
solvant (30)	rozpouštědlo
solvant organique (34)	organické rozpouštědlo
solvant recyclé (2)	recyklované rozpouštědlo

Tabulka č. 31: Kolokace termínu „substance“

Kolokace	Česky - 1
substance (127)	látka
substance chimique (6)	chemická látka
substance radioactive (6)	radioaktivní látka
substance d'origine humain (7)	látka lidského původu
substance dérivée du sang humain (14)	derivát z lidské krve
substance active (7)	účinná látka
stupéfiants et des substances psychotropes (11)	omamné a psychotropní látky
substance protéique (2)	bílkovinná látka

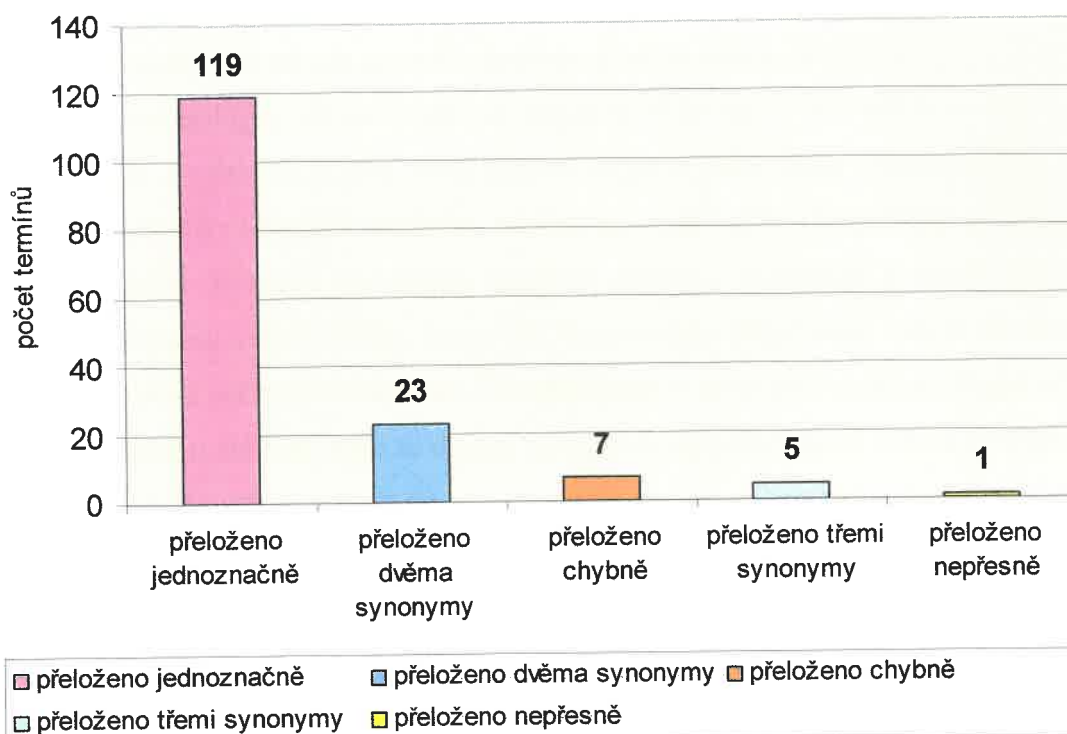
3.4.5 Výsledky analýzy

Pokud bychom chtěli na základě této analýzy shrnout, jak jsou termíny v evropských legislativních textech v oblasti „Ochrana zdraví“ překládány, je třeba si uvědomit, že tato analýza nezkoumala každé slovo, jež je obsaženo v textech, nýbrž jen malou část, tedy přesněji 158 různých termínů či spojení více termínů, tzv. „kolokací“.

Z tohoto celkového počtu termínů či kolokací bylo sedm přeloženo chybně (kapitola 3.4.1) a jeden termín přeložen nepřesně (kapitola 3.4.2). Celkem 23 termínů či kolokací z celkového počtu bylo přeloženo dvěma synonymy, 5 jich pak bylo přeloženo třemi synonymy (kapitola 3.4.3), zbylých 119 termínů či kolokací je přeloženo jednoznačně a souhlasí s použitými slovníky, popřípadě databází ISAP.

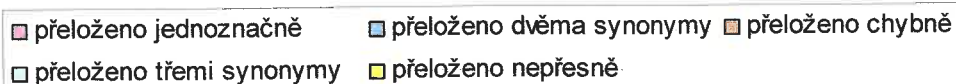
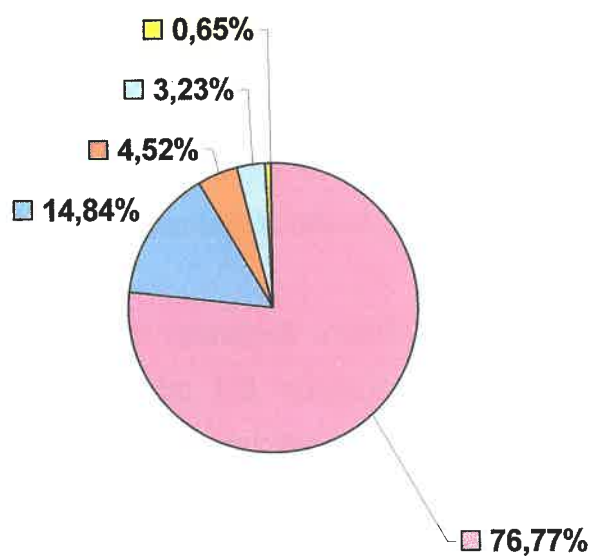
Pro znázornění uvádím dva grafy znázorňující jak překlady termínů v číslech tak i v procentech.

Překlady termínů v číslech



Graf č. 1: Překlady termínů v číslech

Překlady termínů v procentech



Graf č. 2: Překlady termínů v procentech

4. ZÁVĚR

Technický pokrok jde neustále dopředu a to ve všech odvětvích, ať už jsou to informační technologie, zdravotnictví či doprava. Kdo by si jen někdy v minulosti představil, že i výzkum jazyka bude pokrokovější a třeba bude realizován pomocí počítačové techniky? Nejspíš málokdo, možná jen snílci. Ale i lingvistika se nakonec dočkala a může být dnes zkoumána mnohem přesněji, moderněji a navíc získávat výsledky výzkumu velmi rychle. Jde o tzv. korpusovou lingvistiku, což je zkoumání jazyka za pomoci počítačové techniky. Do naší země se tento styl výzkumu dostal až po pádu komunismu, ale i tak jsme se dostali na úroveň západních zemí a dnes se můžeme pyšnit Českým národním korpusem, jež obsahuje na 200 milionů slov českého jazyka. Avšak tímto způsobem nejsou zkoumány jen jednotlivé jazyky ale i více jazyků najednou. Na základě takových výzkumů pak vznikají překladové slovníky, gramatiky a jiné jazykové příručky, umožňují zájemcům studovat jazyk a obohacovat si slovní zásobu netradiční formou. Korpus, jež zkoumá několik jazyků najednou se nazývá paralelní korpus.

Takový projekt vzniká i na katedře romanistiky PF JČU pod vedením Mgr. Jana Radimského, Ph.D. a měl by posloužit všem zájemcům o studium cizího jazyka netradiční formou. Součástí tohoto projektu jsou i texty, které byly předmětem mé analýzy, respektive vybraná klíčová slova z těchto textů. Analýza byla prováděna pomocí programu ParaConc, jež dokáže předem upravené texty paralelně zkoumat a nabízí nám velmi rychlé a přesné výsledky vyhledávání, dále pak i informace o výskytu termínů nebo jejich frekvenci. Podrobnější návod jak s tímto programem pracovat je uveden ve třetí kapitole.

Samotná analýza termínů vybraných z textů z oblasti „Ochrana zdraví“ byla velice zajímavá. Z celkového počtu 158 vybraných termínů nebo spojení několika termínů (tzv. „kolokace“) bylo sedm přeloženo chybně (v podstatě jde o dva chybně přeložené termíny, avšak jsou součástí jiných kolokací, tudíž i ty jsou pak přeloženy nesprávně). Jeden termín byl přeložen nepřesně tzn., že jeho překlad neodpovídal žádnému slovníku, jež jsem měla k dispozici, ani databázi ISAP, avšak vzhledem k tomu, že význam termínu z originálních textů byl zachován, nelze tento překlad označovat jako chybný. Některé termíny jsou v textech překládány více synonymy, 23

termínů je překládáno dvěma synonymy, 5 termínů pak třemi synonymy. 119 termínů je překládáno jednoznačně.

Zjištění, že v textech o 212ti normostranách byly objeveny jen dva opravdu špatně přeložené termíny nám dává pocit, že překládané texty jsou důvěryhodné a pomocí paralelních korpusů mohou být zdrojem pro výuku cizích jazyků moderním a netradičním způsobem, jež snad za pár let bude blízký všem.

5. PŘÍLOHY

5. 1 Slovníček termínů a jejich kolokací

accès de tous aux médicaments; s.m. - všeobecná dostupnost léků, str. 49

accident nucléaire; s.m. - jaderná havárie, str. 57

affection; s.f. - onemocnění, nemoc, str. 45

affection entraînant une menace pour la vie ou une invalidité chronique; s.f. - život ohrožující nebo chronicky invalidizující onemocnění, str. 45

affection mettant la vie en danger, très invalidante ou grave et chronique; s.f. - život ohrožující, závažně invalidizující nebo závažná a chronická onemocnění, str. 45

affection rare; s.f. - vzácné onemocnění, str. 45

Agence européenne pour l'évaluation des médicaments (EMA); s.f. - Evropská agentura pro hodnocení léčivých přípravků (EMA), str. 49

Agence internationale de l'énergie atomique (AIEA); s.f. - Mezinárodní agentura pro atomovou energii (IAEA), str. 56

benzène et le monoxyde de car. dans l'air ambiant; s.m. - benzen a ox. uhelnatý v ovzduší, str. 50

cancer; s.m. - rakovina, str. 55

classification des implants mammaires comme dispositifs médicaux de la classe III; s.f. - klasifikace prsních implantátů jako zdravotnické prostředky třídy III, str. 56

colza; s.m. - řepice, str. 39

colza de printemps; s.m. - jarní řepice, str. 39

colza génétiquement modifié; s.m. - geneticky modifikovaná řepice, str. 39

colza hybride; s.m. - hybridní řepice, str. 39

colza non génétiquement modifié; s.m. - geneticky nemodifikovaná řepice, str. 39

combustible nucléaire; s.m. - jaderné palivo, str. 57

comité; s.m. - výbor, str. 43

Comité "Dispositifs médicaux"; s.m. - Výbor pro zdravotnické prostředky, str. 43

Comité "Normes et règles techniques"; s.m. - Výbor pro normy a technické předpisy, str. 43

comité consultatif; s.m. - poradní výbor, str. 43

comité des médicaments orphelins; s.m. - Výbor pro léčivé přípravky pro vzácná onemocnění, str. 43

Comité des régions; s.m. - Výbor regionů, str. 43

Comité économique et social; s.m. - Hospodářský a sociální výbor, str. 43

Comité europ. de lutte antidrogues (CELAD); s.m. - Evropský výbor pro boj s drogami (CELAD), str. 43

Comité europ. de normalisation (CEN); s.m. - Evropský výbor pro normalizaci (CEN), str. 43

comité scientifique; s.m. - vědecký výbor, str. 43

comité scientifique et technique; s.m. - Výbor pro vědu a techniku, str. 43

Communauté europ. de l'énergie atomique; s.f. - Evropské společenství pro atomovou energii, str. 56

composé; s.m. - sloučenina, látka, str. 46

composé; adj. – složený, str. 46

composé biologique; s.m. - biologická látka, str. 46

composé organique; s.m. - organická sloučenina, str. 46

composé organique volatil (COV); s.m. - těkavá organická sloučenina, str. 46

conséquences du tabagisme sur la santé; s.f. - zdravotní důsledky kouření, str. 53

demande de drogue(s); s.f. - poptávka po drogách, str. 47

dispositif; s.m. – prostředek, str. 56

dispositif destiné à l'évaluation des performances; s.m. - prostředek pro hodnocení funkční způsobilosti, str. 56

dispositif médical; s.m. - zdravotnický prostředek, str. 56

dispositif médical de diagnostic; s.m. - diagnostický zdravotnický prostředek, str. 56

dispositifs destinés à des autodiagnostic; s.m. - prostředky pro sebetestování, str. 56

dissémination; s.f. - uvolňování, rozšiřování, str. 46

dissémination volontaire; s.f. - záměrné uvolňování, str. 46

drogue; s.f. - droga, omamná látka, str. 47

drogues et des toxicomanies; s.f. - drogy a drogová závislost, zneužívání drog, toxikomanie, str. 47

échantillonneur mesurant la pollution; s.m. - měřicí stanice, str. 41

en fonction de l'urg. de la question en cause; s.f. - s ohledem na naléhavost věci, podle naléhavosti věci, str. 54

énergie; s.f. – energie, str. 56

énergie électrique; s.f. - elektrická energie, str. 56

énergie gazeuze; s.f. – plyn, str. 56

énergie hydraulique; s.f. - hydraulická energie, str. 56

énergie pneumatique; s.f. - pneumatická energie, str. 56

épidémie; s.f. – epidemie, str. 57

épidémie du VIH/sida; s.f. - epidemie HIV/AIDS, str. 57

expansion de l'épidémie; s.f. - rozšiřování epidemie, str. 57

fabrication illicite de drogues; s.f. - nedovolená výroba drog, str. 47

gène; s.m. – gen, str. 57

gène bar issu; s.m. - bar-gen, str. 57

gène barnas issu; s.m. - gen pro barnasu, str. 57

gène barstar issu; s.m. - gen barstar, str. 57

gène neo issu; s.m. - gen neo, str. 57

gravité de l'épidémie; s.f. - závažnost epidemie, str. 57

herbicide; s.m. - herbicid, prostředek proti plevelu, str. 48

herbicide chimique; s.m. - chemický herbicid, str. 48

impact de l'épidémie sur...; s.m. - dopad epidemie na..., str. 57

installation nucléaire; s.f. - jaderné zařízení, str. 57

l'Europe contre le cancer; s.f. - Evropa proti rakovině, str. 55

lutte contre la drogue; s.f. - boj s drogami, str. 47

maladie; s.f. - nemoc, onemocnění, str. 37

maladie cardiaque; s.f. - onemocnění srdce, str. 37

maladie cardio-vasculaire; s.f. – onemocnění srdce a cév, str. 37

maladie d'origine alimentaire; s.f. - nemoc přenášená potravou, str. 37

maladie d'origine hydrique et environnementale; s.f. - nemoc přenášená vodou a nemoc závislá na prostředí, str. 37

maladie due à la pauvreté; s.f. - nemoc spojená s chudobou, str. 37

maladie héréditaire; s.f. - dědičné onemocnění, str. 37

maladie rare; s.f. - vzácná nemoc, str. 37

maladie transmissible; s.f. - přenosná nemoc, str. 37

maladie transmissible par des agents non conventionnels; s.f. - nemoc přenášená nekonvenčními činiteli, str. 37

matière nucléaire; s.f. - jaderný materiál, str. 57

médicament; s.m. - léčivý přípravek, lék, léčivo, str. 49

médicament à usage humain; s.m. - humánní léčivý přípravek, str. 49

médicament dérivé du sang ou du plasma humains; s.m. - léčivý přípravek pocházející z lidské krve nebo lidské plazmy, str. 49

médicament orphelin; s.m. - léčivý přípravek pro vzácná onemocnění, str. 49

médicament orphelin autorisé; s.m. - registrovaný léčivý přípravek, str. 49

médicament similaire; s.m. - podobný léčivý přípravek, str. 49

mesure d'urgence; s.f. - havarijní opatření, ochranné opatření, str. 54

mesure d'urgence prévues pour alerter; s.f. - opatření k varování, str. 54

mesure de protection; s.f. - ochranné opatření, str. 58

micro-organisme; s.m. - mikroorganismus, str. 58

ministre de la santé; s.m. - ministr zdravotnictví, str. 52

monoxyde de carbone; s.m. - oxid uhelnatý, str. 50

nucléaire; adj. - jaderný, str. 57

observatoire européen des drogues et des toxicomanies; s.m. - Evropské monitorovací centrum pro drogy, str. 47

organisme; s.m. - organismus, str. 58

organisme 2; s.m. - orgán, subjekt, organizace, str. 51

organisme compétent; s.m. - příslušný subjekt, str. 51

organisme de droit privé; s.m. - soukromoprávní subjekt, str. 51

organisme de protection de l'environnement; s.m. - organizace na ochranu životního prostředí, str. 51

organisme de radiodiffusion; s.m. - rozhlas, str. 51

organisme génétiquement modifié; s.m. - geneticky modifikovaný organismus, str. 58

organisme international; s.m. - mezinárodní orgán, mezinárodní subjekt, str. 51

organisme notifié; s.m. - oznámený subjekt, str. 51

organismes gouvernementaux ou non gouvernementaux; s.m. - vládní a nevládní subjekty, str. 51

phénomène de la drogue; s.m. - problematika drog, jev drog, oblast drog, str. 47

plan d'action contre le cancer; s.m. - akční plán boje proti rakovině, str. 55

plante de tabac; s.f. - odrůda tabáku, tabáková rostlina, str. 53

polluant; adj. – znečišťující, str. 41

polluant; s.m. - znečišťující látka, str. 41

polluant atmosphérique; s.m. - látka znečišťující ovzduší, str. 41

pollution; s.f. - znečištění, znečišťující látka, str. 41

pollution atmosphérique; s.f. - znečišťování ovzduší, str. 41

pollution de l'atmosphère; s.f. - znečištění ovzduší, str. 41

porter atteinte à la santé; v.t. - způsobit zdravotní újmu, str. 52

prévention du cancer; s.f. - prevence rakoviny, str. 55

produits du tabac; s.m. - tabákové výrobky, str. 53

programme d'action contre le cancer; s.m. - akční plán boje proti rakovině, str. 55

protection; s.f. – ochrana, str. 58

protection contre...; s.f. - ochrana před..., str. 58

protection de l'environnement; s.f. - ochrana životního prostředí, str. 58

protection de la santé; s.f. - ochrana zdraví, str. 58

protection sanitaire; s.f. - ochrana zdraví, str. 58

radioprotection; s.f. - radiační ochrana, str. 58

réacteur nucléaire; s.m. - jaderný reaktor, str. 57

résistant à herbicide (bromoxynil); adj. - odolný vůči herbicidu (bromoxynil), str. 48

santé; s.f. - zdraví, zdravotní stav, str. 52

santé d'un patient; s.f. - zdravotní stav pacienta, str. 52

santé de la population; s.f. - zdraví obyvatelstva, str. 52

santé des consommateurs; s.f. - zdraví spotřebitele, str. 52

santé des personnes; s.f. - zdraví jednotlivců, lidské zdraví, str. 52

santé humaine; s.f. - lidské zdraví, str. 52

santé publique; s.f. - veřejné zdraví, lidské zdraví, str. 52

santé sexuelle et génésique; s.f. - sexuální a reprodukční zdraví, str. 52

santé sexuelle et reproductrice; s.f. - sexuální a reprodukční zdraví, str. 52

semences de colza; s.f. - osiva řepice olejné, str. 39

solvant; s.m. – rozpouštědlo, str. 58

solvant organique; s.m. - organické rozpouštědlo, str. 58

solvant recyclé; s.m. - recyklované rozpouštědlo, str. 58

stupéfiants et des substances psychotropes; s.m. - omamné a psychotropní látky, str. 59

substance; s.f. – látka, str. 59

substance active; s.f. - účinná látka, str. 59

substance d'origine humain; s.f. - látka lidského původu, str. 59

substance dérivée du sang humain; s.f. - derivát z lidské krve, str. 59

substance chimique; s.f. - chemická látka, str. 59

substance protéique; s.f. - bílkovinná látka, str. 59

substance radioactive; s.f. - radioaktivní látka, str. 59

tabac; s.m. – tabák, str. 53

tabac à rouler; s.m. - tabák na balení cigaret, str. 53

tabac à usage orale; s.m. - tabák pro orální užití; tabákový výrobek, který se nekouří, str. 53

tabac brut; s.m. - surový tabák, str. 53

tabagisme actif ou passif; s.m. - aktivní nebo pasivní kouření, aktivní nebo pasivní užívání cigaret, str. 53

teneur en monoxyde de carbone des cigarettes; s.f. - úroveň oxidu uhelnatého v cigaretách, obsah oxidu uhelnatého v cigaretách, míra ox. uhelnatého v cigaretách, str. 50

tolérant aux herbicides; adj. - tolerantní k herbicidům, str. 48

transfert du gène de résistance à...; s.m. - přenos genu tolerance k..., str. 57

transmission du VIH/sida; s.f. - přenos HIV/AIDS, str. 42

urgence; s.f. - opatření, mimořádná situace, str. 54

urgence radiologique; s.f. - radiační mimořádná situace, str. 54

VIH; s.m. – HIV, str. 42

VIH/sida; s.m. - HIV/AIDS, str. 42

5. 2 Směrnice Komise upravující názvy „řepka“ a „řepice“

Směrnice Komise

ze dne 27. června 1979,

kterou se mění směrnice Rady 66/401/EHS, 66/402/EHS, 69/208/EHS a 70/458/EHS o uvádění osiva pícein, osiva obilovin, osiva olejnin a přadných rostlin a osiva zeleniny na trh

(79/641/EHS)

KOMISE EVROPSKÝCH SPOLEČENSTVÍ,

s ohledem na Smlouvu o založení Evropského hospodářského společenství,

s ohledem na směrnici Rady 66/401/EHS ze dne 14. června 1966 o uvádění osiva pícein na trh, naposledy pozměněnou směrnicí 78/1020/EHS, a zejména na čl. 2 odst. 1a a článek 21a uvedené směrnice,

s ohledem na směrnici Rady 66/402/EHS ze dne 14. června 1966 o uvádění osiva obilovin na trh, naposledy pozměněnou směrnicí 78/1020/EHS, a zejména na čl. 2 odst. 1a a článek 21a uvedené směrnice,

s ohledem na směrnici Rady 69/208/EHS ze dne 30. června 1969 o uvádění osiva olejnin a přadných rostlin na trh, naposledy pozměněnou směrnicí 78/1020/EHS, a zejména na čl. 2 odst. 1a a článek 40a uvedené směrnice,

s ohledem na směrnici Rady 70/458/EHS ze dne 29. září 1970 o uvádění osiva zeleniny na trh, naposledy pozměněnou směrnicí 78/692/EHS, a zejména na čl. 2 odst. 1a a článek 40a uvedené směrnice,

vzhledem k tomu, že se v důsledku vývoje vědeckých a technických poznatků ukázalo, že některé botanické názvy použité ve směrnicích o uvádění osiv na trh jsou nesprávné nebo jejich hodnověrnost je nejistá;

vzhledem k tomu, že by se tyto názvy měly sjednotit s názvy běžně užívanými na mezinárodním úrovni;

vzhledem k tomu, že by proto měly být změněny směrnice 66/401/EHS, 66/402/EHS, 69/208/EHS a 70/458/EHS o uvádění osiva pícein, osiva obilovin, osiva olejnin a přadných rostlin a osiva zeleniny na trh;

vzhledem k tomu, že některá ustanovení uvedených směrnic a směrnice Rady 70/457/EHS ze dne 29. září 1970 o Společném katalogu odrod dřevěných rostlin...

rostlin, naposledy pozměněné směrnicí 78/55/EHS, používají pojem druh, čímž určují oblast působnosti uvedených ustanovení; že změny nomenklatury podle této směrnice by neměly ovlivnit výklad těchto ustanovení;

vzhledem k tomu, že opatření této směrnice jsou v souladu se stanoviskem Stálého výboru pro osivo a sadbu v zemědělství, zahradnictví a lesnictví,

PŘIJALA TUTO SMĚRNICI:

Článek 1

Směrnice Rady 66/401/EHS o uvádění osiva píce na trh se mění takto:

1. V čl. 2 odst. 1 části A písm. a) se názvy druhů:

Agrostis canina L. ssp. *canina* Hwd. – psineček psí,

Arrhenatherum elatius (L.) J. et. C. Presl. – ovsík vyvýšený,

Phleum bertolinii DC – bojínek cibulkatý,

Trisetum flavescens (L.) Pal. Beauv. – trojštět žlutavý,

nahrazují tímto:

Agrostis canina L. – psineček psí,

Arrhenatherum elatius (L.) Beauv. ex J. et K. Presl. – ovsík vyvýšený,

Phleum bertolonii DC – bojínek cibulkatý,

Trisetum flavescens (L.) Beauv. – trojštět žlutavý.

2. V čl. 2 odst. 1 části A písm. b) se názvy druhů:

Medicago varia Martyn – vojtěška proměnlivá,

Onobrychis sativa Lam. – vičenec setý,

Pisum arvense L. – hrách krmný,

Trigonella foenumgraecum L. – pískavice řecké seno,

nahrazují tímto:

Medicago x varia Martyn – vojtěška proměnlivá,

Onobrychis viciifolia Scop. – vičenec setý

Pisum sativum L. (partim) – hrách setý, krmný,

Trigonella foenum-graecum L. – pískavice řecké seno.

3. V čl. 2 odst. 1 části A písm. b) se názvy druhů:

Vicia faba L. ssp. *faba* var. *equina* Pers. – bob polní,

Vicia faba L. var. *minor* (Peterm.) Bull – bob polní menší,

nahrazují tímto:

Vicia faba L. (partim) – bob.

4. V čl. 3 odst. 1 se názvy druhů:

Medicago varia Martyn,

Pisum arvense L.

nahrazují tímto:

Medicago x varia Martyn,

Pisum sativum L.

5. V příloze I bodech 2 a 4 a v příloze II části I bodě 1 a části II bodě 1 název druhu:

Pisum arvense,

nahrazuje tímto:

Pisum sativum.

6. V příloze II části I bodě 2 oddíle A a části II bodě 2 oddíle A a v příloze III se názvy druhů:

Agrostis canina ssp. *canina*,

Medicago varia,

Onobrychis sativa,

Pisum arvense,

nahrazují tímto:

Agrostis canina,

Medicago x varia,

Onobrychis viciifolia,

Pisum sativum.

7. V příloze II části I bodě 2 oddíle A a části II bodě 2 oddíle A a v příloze III se názvy druhů:

Vicia faba ssp. var. *equina*,

Vicia faba var. *minor*,

nahrazují tímto:

Vicia faba.

8. V příloze II části I bodě 2 oddíle B písm. e) se názvy druhů:

Pisum arvense,

Vicia faba spp.,

nahrazují tímto:

Pisum sativum,

Vicia faba.

Článek 2

Směrnice Rady 66/402/EHS o uvádění osiva obilovin na trh se mění takto:

1. V čl. 2 odst. 1 části A se názvy druhů:

Hordeum distichum L. – ječmen dvouřadý,

Hordeum polystichum L. – ječmen šestiřadý,

nahrazují tímto:

Hordeum vulgare L. – ječmen.

2. V čl. 2 odst. 1 části A se názvy druhů:

Triticum aestivum L. – pšenice obecná,

Triticum durum L. – pšenice tvrdá,

nahrazují tímto:

Triticum aestivum L. emend. Fiori et Paol. – pšenice obecná

Triticum durum Desf. – pšenice tvrdá.

3. V příloze I bodě 5 oddíle B písm. a), v příloze II bodě 1 oddíle A a bodě 2 oddíle A a v příloze III se názvy druhů:

Hordeum distichum,

Hordeum polystichum,

nahrazují tímto:

Hordeum vulgare.

Článek 3

Směrnice Rady 69/208/EHS o uvádění osiva olejnin a přadných rostlin na trh se mění takto:

1. Názvy druhů v čl. 2 odst. 1 části A se nahrazují tímto:

Arachis hypogaea L. – podzemnice olejná,

Brassica juncea (L.) Czern. et Coss. in Czern. – hořčice sareptská,

Brassica napus L. ssp. *oleifera* (Metzg.) Sinsk. – řepka olejka,

Brassica nigra (L.) W. Koch – hořčice černá,

Brassica rapa L. (partim) – řepice olejná,

Cannabis sativa L. – konopí seté,

Carum carvi L. – kmín kořený,

Glycine max (L.) Merr. – soja luštinatá,

Gossypium spp. – bavlník,

Helianthus annuus L. – slunečnice roční,

Linum usitatissimum L. – len setý,

Papaver somniferum L. – mák setý,

Sinapis alba L. – hořčice bílá.

2. V čl. 3 odst. 1 se zrušuje název druhu *Brassica campestris* L. ssp. *oleifera* (Metzg.) Sinsk.

Za název *Brassica napus* L. ssp. *oleifera* (Metzg.) Sinsk. se vkládá název *Brassica rapa* L. (partim).

3. V příloze I bodě 2 se název druhu *Brassica napus oleifera* nahrazuje názvem *Brassica napus ssp. oleifera*.

4. V příloze II části I bodě 3 oddíle A se zrušuje název druhu *Brassica campestris ssp. oleifera*.

Za název *Brassica napus ssp. oleifera* se vkládá název *Brassica rapa*.

5. V příloze III se název druhu *Brassica campestris spp. oleifera* nahrazuje názvem *Brassica rapa*.

Článek 4

Směrnice Rady 70/458/EHS o uvádění osiva zeleniny na trh se mění takto:

1. V čl. 2 odst. 1 části A se názvy druhů:

Citrullus vulgaris L. – meloun vodní,

Daucus carota L. ssp. *sativus* (Hoffm.) Hayek – mrkev obecná,

Petroselinum hortense Hoffm. – petržel kadeřavá,

Pisum sativum L. (excl. *P. arvense* L.) – hrách setý,

Valerianella locusta (L.) Betcke (*V. olitoria* Polt.) – kozlíček polníček,

Vicia faba major L. – bob,

nahrazují tímto:

Citrullus lanatus (Thumb.) Matsum. et Nakai – meloun vodní,

Daucus carota L. – mrkev obecná,

Petroselinum crispum (Mill.) Nym. ex A. W. Hill – petržel kadeřavá,

Pisum sativum L. (partim) – hrách setý, vyjma krmného,

Valerianella locusta (L.) Laterr. – kozlíček polníček,

Vicia faba L. (partim) – bob.

2. V čl. 2 odst. 1 část A se zrušuje název druhu *Solanum lycopersicum* L. (*Lycopersicum esculentum* Mill.) – rajče.

Za název *Lactuca sativa* L. – salát se vkládá název *Lycopersicon lycopersicum* (L.) Karst. ex Farwell – rajče.

3. V příloze II bodě 3 písm. a) a v příloze III bodě 2 se názvy druhů:

Citrullus vulgaris,

Petroselinum hortense,

Solanum lycopersicum,

nahrazují tímto:

Citrullus lanatus,

Petroselinum crispum,

Lycopersicon lycopersicum.

Článek 5

Touto směrnicí není dotčeno použití ostatních ustanovení směrnic 66/401/EHS, 66/402/EHS, 69/208/EHS a 70/458/EHS ani ustanovení směrnice 70/457/EHS.

Článek 6

Členské státy uvedou v účinnost právní a správní předpisy nezbytné pro dosažení souladu s ustanoveními této směrnice nejpozději 1. července 1980. Neprodleně o nich uvědomí Komisi.

Článek 7

Tato směrnice je určena členskými státy.

V Bruselu dne 27. června 1979.

Za Komisi

Finn Gundelach

místopředseda

(<http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=CELEX:31979L0641:CS:HTML>)
[staženo 15. 4. 2006]

PRAMENY A LITERATURA

Odborné publikace a články:

BLATNÁ, Renata. Korpus - uskutečněný sen. *Tvar*. 1997, č. 9, str.17. ISSN 0862-657X

ČERMÁK, F. Jazykový korpus: Prostředek a zdroj poznání. *Slovo a Slovesnost*. 1995, č. 56, str. 119 – 140. ISSN 0037-7031

ČERMÁK, F.; KUBÍČEK, P. Jazykový korpus a škola. *Český jazyk a literatura*. 1997-98, roč. 48, č. 3-4, str. 84-92

KOCEK, Jan; KOPŘIVOVÁ, Marie; KUČERA, Karel. *Český národní korpus: Úvod a příručka uživatele*. 1. vyd. Praha: UK v Praze - ÚČNK, 2000. 156 s. ISBN 80-85899-94-9. Část dostupná na URL <<http://knihovna.nkp.cz/nkkr0403/0403152.html>>

ŠTÍCHA, F. Čas korpusové lingvistiky. *Slovo a slovesnost*. 1994, č. 55, str. 141 – 145. ISSN 0037-7031

ŠULC, Michal. *Korpusová lingvistika: První vstup*. 1. vyd. Praha: UK v Praze - nakl. Karolinum, 1999. 94 s. ISBN 80-7184-847-6

Slovníky:

Collectif. *Larousse de poche*. Paris: Larousse, 2004. 1008 s. ISBN 2-03-532085-2

KRAUS Jiří; kolektiv autorů. *Nový akademický slovník cizích slov A – Ž*. 1. vyd. Praha: Academia, 2005. 880 s. ISBN 80-200-1351-2

Lingea s.r.o. *Francouzsko-český všeobecný slovník* [CD ROM]. Verze 4.11. Lingea Lexicon, 2002

NEUMANN, Josef; HOŘEJŠÍ, Vladimír; kolektiv. *Velký francouzsko-český slovník I. (A – K)*. 2. vyd. Praha: Academia, 1992. 833 s. ISBN 80-200-0234-0

NEUMANN, Josef; HOŘEJŠÍ, Vladimír; kolektiv. *Velký francouzsko-český slovník II. (L – Z)*. 2. vyd. Praha: Academia, 1992. 757 s. ISBN 80-200-0235-9

Elektronické zdroje:

BLATNÁ, Renata. Jazyková databanka neboli korpus. *Vesmír*. 1997, č. 12, str. 670-671. Dostupný na <<http://www.vesmir.cz/clanek.php3?CID=3118>> [cit. 3. 4. 2006]. ISSN 1214-4029

ČERMÁK, F. Dvoujazyčný slovník a korpus: některé poznámky. Článek dostupný na <ucnk.ff.cuni.cz/doc/2002_cnk.rtf> [staženo 6. 4. 2006].

ČERMÁK, F. Korpus, informace a lingvistika. Článek dostupný na <ucnk.ff.cuni.cz/doc/korp-info-lingv.rtf> [staženo 6. 4. 2006].

KOPŘIVOVÁ, M. Český národní korpus na přelomu tisíciletí. Článek dostupný na <ucnk.ff.cuni.cz/doc/korplex.rtf> [staženo 6. 4. 2006].

SCHMIEDTOVÁ, Věra. Český národní korpus. *Ikaros* [online]. 1999, roč. 3, č. 5. Dostupný na <<http://www.ikaros.cz/node/355>> [cit. 3. 4. 2006]. ISSN 1212-5075

<http://www.vesmir.cz/>

<http://ucnk.ff.cuni.cz/>

<http://ucnk.ff.cuni.cz/korpus/korpus.html>

<https://trnka.ff.cuni.cz/ucnk/intercorp/>

<http://www.athel.com/para.html>

<http://europa.eu.int/eur-lex/>

<http://isap.vlada.cz/>

<http://genbank.vurv.cz/>

POUŽITÉ ZKRATKY

s – substantiv

f – féminin

m – masculin

v – verbe

t – transitif

adj – adjectif

RÉSUMÉ (en français)

Premièrement il faut se poser une question qu'est-ce que c'est un corpus? Dans la domaine de linguistique le corpus est une un ensemble structuré de textes organisées selon des critères linguistiques et élaborés à l'aide de l'informatique. Le but du corpus est d'être représentatif au maximum d'une langue et d'une variété de langue et d'examiner la langue d'une manière peu traditionnelle. Il peut servir à développer des dictionnaires, des grammaires ou des manuels linguistiques et à étudier une langue par le façon insolite. La linguistique de corpus permet d'observer plus précisément la variation de la langue.

L'utilisation de corpus de texte électroniques existe depuis les années soixante. Dans la République tchèque la linguistique de corpus ne cesse de croître depuis 1994. Aujourd'hui il existe deux corpus écrits de la langue tchèque (SYN2000 et SYN2005) possédant ensemble 200 millions de mots. La plupart des langues de la communauté européenne possèdent un ou plusieurs corpus électroniques (par exemple en France il y a un grand corpus de la langue française appelé Frantext).

Le but de mon travail est d'analyser la terminologie française et tchèque dans les textes législatifs de la communauté européenne. Cette analyse était faite à l'aide d'un logiciel appelé ParaConc qui est capable d'examiner deux ou plusieurs langues ensemble (quatre langues au maximum). Le ParaConc ne nous offre pas seulement la recherche rapide des mots dans plusieurs langues mais aussi des informations des fréquences des mots ou leurs collocations (des mots souvent liés avec le mot clé).

Mon analyse peut nous montrer comment les traductions des textes législatifs sont précises. Du nombre total des textes analysés (158), deux mots ne sont pas traduits correctement, un mot est traduit inexactement, ça veut dire que la traduction ne correspond pas aux dictionnaires mais le sens des phrases comprenant ce mot restent sans changement. 23 mots sont traduits par deux synonymes, 5 mots par trois synonymes. Le reste des mots est traduit nettement et sans ambiguïté. D'après cette analyse on peut constater que dans les traductions des textes législatifs de l'UE on ne trouve pas les grandes erreurs et ils peuvent nous servir comme un matériel d'enseignement.

Anotace

Tato bakalářská práce popisuje v první řadě relativně nový jazykový obor, což je korpusová lingvistika. Nastiňuje její historii, vysvětluje pojem „korpus“ a popisuje druhy jazykových korpusů. Dále se zabývá i Českým národním korpusem, jeho historií a přínosem pro studování českého jazyka. Blíže se pak tato práce zabývá paralelním korpusem, což je korpus, ve kterém se zkoumají dva i více cizojazyčných textů dohromady. Popisuje práci s programem ParaConc pomocí něhož se cizojazyčné texty paralelně zkoumají. Výzkumná část této práce má za úkol analyzovat a zhodnotit překlady vybraných termínů v evropských legislativních textech zaměřených na oblast „Ochrana zdraví“.

Annotation

Primarily this thesis describes a new branch in the linguistics – the Corpus Linguistics. The thesis deals with the history of corpus linguistics and it explains the term “corpus” and describes the types of corpus. It deals also with the Czech National Corpus its history and its contribution to the research on the Czech language. Next, this thesis deals with parallel corpus which can examine more languages together. It describes how to work with the program called “ParaConc” which allow to analyze the parallel texts. The research target of this thesis is to analyze and evaluate the translations of terms in the European legislative texts from the domain: „The health protection“.