

JIHOČESKÁ UNIVERZITA V ČESKÝCH BUDĚJOVICÍCH

FILOZOFICKÁ FAKULTA

ÚSTAV ROMANISTIKY

BAKALÁŘSKÁ PRÁCE

FRANCOUZSKÁ A ČESKÁ TERMINOLOGIE V EVROPSKÝCH  
LEGISLATIVNÍCH TEXTECH ZAMĚŘENÝCH NA OBLAST: „VYBRANÉ  
ZEMĚDĚLSKÉ VÝROBKY – TABÁK, LEN, KONOPÍ, CHMEL“

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

Autorka práce: Lenka Hříbalová

Studijní program: Filologie

Studijní obor: Francouzský jazyk pro evropský a mezinárodní obchod

Ročník: 3.

2010

***Bakalářská práce***

Jihočeská univerzita v Českých Budějovicích

Filozofická fakulta

Ústav romanistiky

**Název:** Francouzská a česká terminologie v evropských legislativních textech zaměřených na oblast: „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“

**Name:** French and Czech terminology in European legislative texts focused on the field: „Selected agricultural products – tobacco, flax, hemp, hop“

**Autorka:** Lenka Hříbalová

**Studijní program:** Filologie

**Studijní obor:** Francouzský jazyk pro evropský a mezinárodní obchod

**Ročník:** 3.

**Vedoucí práce:** PhDr. Jan Radimský, Ph.D.

České Budějovice 2010

### **Prohlášení**

Prohlašuji, že jsem bakalářskou práci na téma Francouzská a česká terminologie v evropských legislativních textech zaměřených na oblast: „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“ zpracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47 zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejich internetových stránkách.

České Budějovice, 20. května 2010

.....

## **Poděkování**

Na tomto místě bych velmi ráda poděkovala PhDr. Janu Radimskému, Ph.D. za jeho pomoc, cenné rady, trpělivost a ochotu při vedení bakalářské práce.

## ANOTACE

Tato bakalářská práce se zabývá francouzskou a českou terminologií evropských legislativních textů zaměřených na oblast „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“.

Práce je rozdělena do dvou částí – teoretické a výzkumné. První část práce, teoretická, se zabývá vybranými lingvistickými pojmy, které se v práci vyskytují a dále vysvětlením pojmů korpus, korpusová lingvistika. Následují vybrané texty, postup přípravy a úpravy textů. Dále představení programu ParaConc, který se zabývá prací s paralelními texty a postup práce s programem. Druhá část práce, výzkumná, se zabývá samostatným výběrem textů, vybraných termínů a kolokací, následované podrobným zkoumáním překladů.

Závěrečná část práce je francouzsko-český slovníček zkoumaných termínů a kolokací.

## ANNOTATION

This bachelor thesis deals with the French and Czech terminology of the European legislative texts aimed at the section „Selected agricultural products – tobacco, flax, hemp, hops.

The work is divided into two parts – theoretical and research. The first part, theoretical, deals with the selection of linguistic terms which occur in the present work and explaining of terms “corpus”, “corpus linguistics”. Then there are selected texts, the process of preparation and editing of texts. Furthermore is presentation of the software ParaConc which deals with parallel texts and the work operation with this software. The second part of the thesis, research, is based on the individual choice of texts, selected terms and collocations continued by the detailed examination of the translations.

The final output of thesis is French-Czech glossary of examined terms and collocations.



## **OBSAH:**

Anotace .....	5
Annotation .....	6
Obsah .....	8
<b>Úvod</b> .....	11
<b>1. Výběr a příprava textů</b> .....	12
1.1. Úvod .....	12
1.2. Příprava textů .....	12
1.2.1. Texty z oblasti „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“ .....	12
1.3. Typy dokumentů - druhy právních předpisů .....	16
<b>2. Zpracování textů, program ParaConc</b> .....	18
2.1. Úprava formátu legislativních textů .....	18
2.2. Vytvoření paralelního korpusu .....	18
2.3. Výběr klíčových slov a kolokací .....	20
2.3.1. Definice pojmů .....	22
2.3.1.1. Pojem kolokace .....	23
2.3.1.1.1. Typy kolokací .....	23
<b>3. Korpus a korpusová lingvistika</b> .....	28
3.1. Korpusová lingvistika .....	28
3.2. Korpus .....	28
3.3. Typy korpusů .....	28
3.3.1. Textový korpus .....	29
3.3.2. Paralelní korpus .....	29
3.3.3. Synchronní korpus .....	30
3.3.4. Diachronní korpus .....	30
3.3.5. Mluvený korpus .....	30
3.3.5.1. Cizí mluvené korpusy .....	31



3.3.5.2. České mluvené korpusy .....	32
3.4. Využití korpusů .....	32
<b>4. Práce s texty .....</b>	<b>33</b>
4.1. Morfologické struktury kolokací .....	33
4.2. Struktury kolokací pro vybrané texty z oblasti: „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“ .....	33
4.3. Závěr .....	37
<b>5. Překlady .....</b>	<b>39</b>
5.1. Adekvátní překlady .....	39
5.1.1. Kolokace s jedním možným překladem .....	39
5.1.2. Kolokace s více synonymními překlady .....	50
5.2. Neadekvátní překlady .....	58
5.3. Závěr výzkumu .....	62
<b>6. Závěr .....</b>	<b>64</b>
<b>7. Resumé .....</b>	<b>65</b>
<b>8. Seznam použité literatury .....</b>	<b>67</b>
<b>9. Přílohy .....</b>	<b>69</b>



# Úvod

Hlavním cílem této bakalářské práce je analyzovat termíny a kolokace francouzských a českých evropských legislativních textů zaměřených na oblast: „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“. Termíny a kolokace budu vybírat ze stránek EUR-Lex, kde jsou dostupné evropské legislativní texty.

V první části bakalářské práce nejprve popíši postup přípravy a úpravy legislativních textů a jejich následné zpracování v programu ParaConc. Poté popíši uložení klíčových slov a kolokací do databáze MS Acces. V další kapitole vysvětlím pojmy korpus a korpusová lingvistika, definuji pojmy – termín, kolokace, klíčové slovo. V následné části se zaměřím na struktury kolokací z morfologického hlediska.

V druhé, hlavní části bakalářské práce popíši práci s vybranými klíčovými slovy a kolokacemi. S pomocí terminologických databází a slovníků budu hodnotit a zkoumat správnost českých překladů těchto klíčových slov a kolokací.

Výsledkem této bakalářské práce bude analýza správnosti českých překladů a následné vytvoření francouzsko-českého slovníčku všech odborných termínů a kolokací, které se vztahují k oblasti „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“.

# 1. Výběr a příprava textů

## 1.1. Úvod

Nejprve zde popíši hledání textů, postup při výběru legislativních textů, ze kterých budou následně vycházet mnou vybraná klíčová slova a kolokace. V této části budu také analyzovat vybrané texty, termíny.

## 1.2. Příprava textů

### 1.2.1. Texty z oblasti „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“

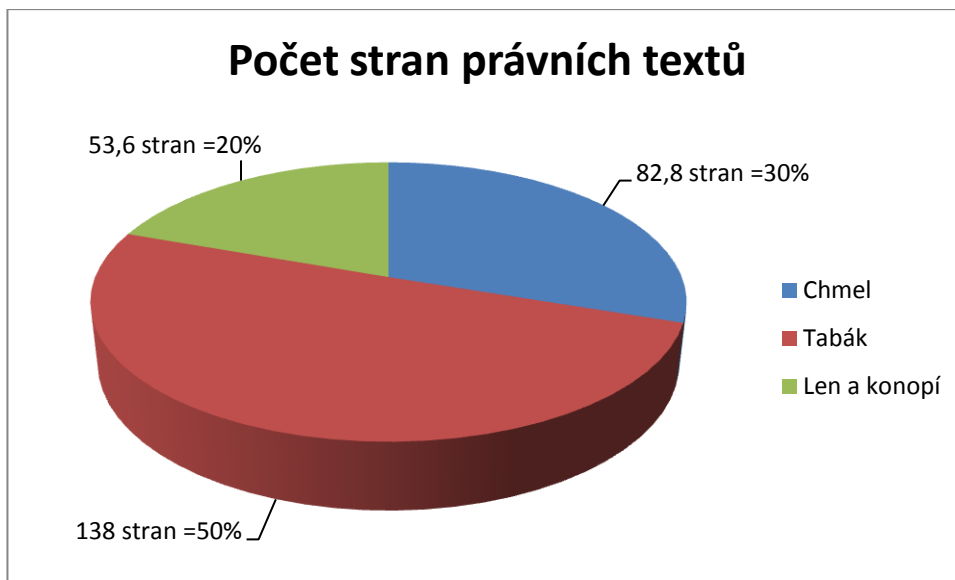
Evropské legislativní texty pro mou oblast „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“ jsou přístupné na stránkách databáze EUR-Lex<sup>1</sup>. Kde stránky nabízejí „základní vyhledávání“, „pokročilé vyhledávání“ a vyhledávání podle „úředního věstníku“. Já jsem pro vyhledávání textů používala přístup ze „základního vyhledávání“, kde jsem si vybírala texty dle „kapitol rejstříku“. Přestože je oblast „kapitoly rejstříku“ obsáhlá a nabízí výběr z 20 různých kapitol, já pro svou práci potřebovala jen kapitolu „Zemědělství“ (03. Agriculture). Z této oblasti jsem si vybrala podskupinu „Produkty, které jsou předmětem společné organizace trhů“ (03.60. Produits faisant l'objet d'une organisation commune des marchés) a z ní jsem si následně vybrala podskupiny „Tabák“ (03.60.64. Tabac), „Len a konopí“ (03.60.65. Lin et chanvre) a „Chmel“ (03.60.66. Houblon). Podskupiny „tabák“, „len a konopí“ a „chmel“ obsahují dohromady 948 legislativních textů, z nichž jsem si pro svou práci vybrala celkem 30 textů z daných oblastí. A to v tomto počtu: 13 textů z oblasti chmele, 12 textů z oblasti tabáku a zbylých 5 textů z oblasti lnu a konopí. Neméně důležitým kritériem byl stanovený počet normostran, a to na 250 stran. V mém případě těch 30 textů dává dohromady 274,4 normostran v českém jazyce a 316,3 normostran v jazyce francouzském. 13 textů z oblasti chmele čítá dohromady 82,8 stran v jazyce českém a 100,7 stran v jazyce francouzském, texty z oblasti tabák čítají 138 stran v jazyce českém a 150,7 stran v jazyce francouzském a texty pro oblast len a konopí čítají 53,6 stran v jazyce českém a 64,9 stran v jazyce francouzském. Pro znázornění jsem vytvořila

---

<sup>1</sup> EUR-Lex dostupný na <http://eur-lex.europa.eu/cs/index.htm> [cit. 11.3.2010].

výšečový graf, který obsahuje počet stran daných textů v českém jazyce a k tomu odpovídající procentuální zastoupení.

Graf č. 1: Počet stran textů



Tyto texty obsahují 70 473 slov v českém jazyce a 90 394 slov v jazyce francouzském. Vybírala jsem dokumenty vydané od roku 1971 do roku 2009.

Pro větší přehlednost jsem vytvořila tabulky, které obsahují důležité informace o textech. Tabulky obsahují CELEXové číslo, číslo a rok vydání v Úředním věstníku (pokud dostupné), rok vydání, počet stran, typ dokumentu a nezkrácený název dokumentu podle databáze EUR-Lex.

Tabulka č. 1: Seznam dokumentů - chmel

### 03.60.66. Chmel

CHMEL					
Celex číslo	Úřední věstník	Rok	Počet stran	Typ dokumentu	Nezkrácený název dokumentu podle EUR-Lex
31992R3124	L 313 1992	1992	2,4	Nařízení	Nařízení Rady (EHS) č. 3124/92 ze dne 26. října 1992, kterým se mění nařízení (EHS) č. 1696/71 o společné organizaci trhu s chmelem
31991R2264	L 208 1991	1991	2,4	Nařízení	Nařízení Komise (EHS) č. 2264/91 ze dne 26. července 1991, kterým se mění nařízení (EHS) č. 3076/78 o dovozu chmele ze třetích zemí

CHMEL					
Celex číslo	Úřední věstník	Rok	Počet stran	Typ dokumentu	Nezkrácený název dokumentu podle EUR-Lex
31988R4060	L 356 1988	1988	2,9	Nařízení	Nařízení Komise (EHS) č. 4060/88 ze dne 21. prosince 1988, kterým se mění některá nařízení v odvětví chmele po zavedení kombinované nomenklatury
31977R1784	L 200 1977	1977	4,1	Nařízení	Nařízení Rady (EHS) č. 1784/77 ze dne 19. července 1977 o vydávání ověřovacích listin původu pro chmel
31993R2928	L 265 1993	1993	4,2	Nařízení	Nařízení Komise (EHS) č. 2928/93 ze dne 25. října 1993, kterým se mění nařízení (EHS) č. 890/78, kterým se stanoví prováděcí pravidla pro vydávání ověřovacích listin původu pro chmel
31991R2265	L 208 1991	1991	4,2	Nařízení	Nařízení Komise (EHS) č. 2265/91 ze dne 26. července 1991, kterým se mění nařízení (EHS) č. 890/78, kterým se stanoví prováděcí pravidla pro vydávání ověřovacích listin původu pro chmel
31999R0609	L 75 1999	1999	6,5	Nařízení	Nařízení Komise (ES) č. 609/1999 ze dne 19. března 1999 o poskytování podpory pro producenty chmele
31978R3076	L 367 1978	1978	6,5	Nařízení	Nařízení Komise (EHS) č. 3076/78 ze dne 21. prosince 1978 o dovozu chmele ze třetích zemí
31997R1554	L 208 1997	1997	7,8	Nařízení	Nařízení Rady (ES) č. 1554/97 ze dne 22. července 1997, kterým se mění nařízení (EHS) č. 1696/71 o společné organizaci trhu s chmelem
31977R1170	L 137 1977	1977	8,1	Nařízení	Nařízení Rady (EHS) č. 1170/77 ze dne 17. května 1977, kterým se mění nařízení (EHS) č. 1696/71 o společné organizaci trhu s chmelem
31978R0890	L 117 1978	1978	9,1	Nařízení	Nařízení Komise (EHS) č. 890/78 ze dne 28. dubna 1978, kterým se stanoví prováděcí pravidla pro vydávání ověřovacích listin původu pro chmel
52005PC0386	KOM/2005/0386	2005	12	Návrh nařízení	Návrh nařízení Rady o společné organizaci trhu s chmelem
31971R1696	L 175 1971	1971	12,6	Nařízení	Nařízení Rady (EHS) č. 1696/71 ze dne 26. července 1971 o společné organizaci trhu s chmelem
			Celkem:82,8		

Tabulka č. 2: Seznam dokumentů - tabák

### 03.60.64. Tabák

## TABÁK

<b>Celex číslo</b>	<b>Úřední věstník</b>	<b>Rok</b>	<b>Počet stran</b>	<b>Typ dokumentu</b>	<b>Nezkrácený název dokumentu podle EUR-Lex</b>
31995R0711	L 73 1995	1995	2,7	Nářízení	Nářízení Rady (ES) č. 711/95 ze dne 27. března 1995, kterým se mění nařízení (EHS) č. 2075/92 o společné organizaci trhu se surovým tabákem
32003L0117	L 333 2003	2003	2,8	Směrnice	Směrnice Rady 2003/117/ES ze dne 5. prosince 2003, kterou se mění směrnice 92/79/EHS a 92/80/EHS, s cílem povolit Francouzské republice prodloužené uplatňování nižších sazeb spotřební daně na tabákové výrobky uvolněné ke spotřebě na Korsice
31992L0080	L 316 1992	1992	3,3	Směrnice	Směrnice Rady 92/80/EHS ze dne 19. října 1992 o sblížení daní z jiných tabákových výrobků než cigaret
32005R1881	L 301 2005	2005	4,3	Nářízení	Nářízení Komise (ES) č. 1881/2005 ze dne 17. listopadu 2005, kterým se mění nařízení (ES) č. 2182/2002, kterým se stanoví prováděcí pravidla k nařízení Rady (EHS) č. 2075/92, pokud jde o Tabákový fond Společenství
31993R0085	L 12 1993	1993	7,6	Nářízení	Nářízení Komise (EHS) č. 85/93 ze dne 19. ledna 1993 o kontrolních subjektech v odvětví tabáku
32008R0709	L 197 2008	2008	8,4	Nářízení	Nářízení Komise (ES) č. 709/2008 ze dne 24. července 2008, kterým se stanoví prováděcí pravidla k nařízení Rady (ES) č. 1234/2007, pokud jde o mezioborové organizace a dohody v odvětví tabáku
31998R1636	L 210 1998	1998	8,6	Nářízení	Nářízení Rady (ES) č. 1636/98 ze dne 20. července 1998, kterým se mění nařízení (EHS) č. 2075/92 o společné organizaci trhu se surovým tabákem
31992R2077	L 215 1992	1992	8,7	Nářízení	Nářízení Rady (ES) č. 2077/92 ze dne 30. června 1992 o mezioborových organizacích a dohodách v odvětví tabáku
31992R2075	L 215 1992	1992	12,2	Nářízení	Nářízení Rady (EHS) č. 2075/92 ze dne 30. června 1992 o společné organizaci trhu se surovým tabákem
31979R1979	L 228 1979	1979	15,9	Nářízení	Nářízení komise (EHS) č. 1979/79 ze dne 5. září 1979, kterým se mění nařízení (EHS) č. 1727/70, (EHS) č. 1728/70, (EHS) č. 2603/71, (EHS) č. 638/74 a (EHS) č. 410/76, pokud jde o jejich použití pro odrůdy tabáku z produkce Společenství
52009AE0635	C 228 2009	2009	19,1	Stanovisko	Stanovisko Evropského hospodářského a sociálního výboru k návrhu směrnice Rady, kterou se mění směrnice 92/79/EHS, 92/80/EHS a 95/59/ES, pokud jde o strukturu a sazby spotřební daně z tabákových výrobků
31998R2848	L 358 1998	1998	44,4	Nářízení	Nářízení Komise (ES) č. 2848/98 ze dne 22. prosince 1998, kterým se stanoví prováděcí pravidla k nařízení (EHS) č.

TABÁK					
Celex číslo	Úřední věstník	Rok	Počet stran	Typ dokumentu	Nezkrácený název dokumentu podle EUR-Lex
					2075/92, pokud jde o režim prémie, produkční kvóty a zvláštní podporu pro seskupení producentů v odvětví surového tabáku
			Celkem:138		

Tabulka č. 3: Seznam dokumentů – len a konopí

### 03.60.65. Len a konopí

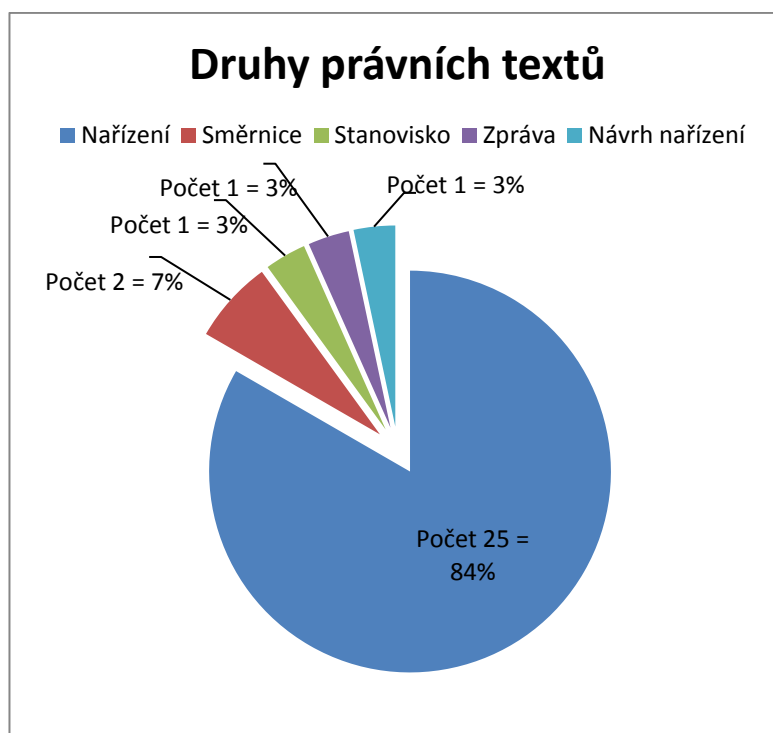
LEN A KONOPÍ					
Celex číslo	Úřední věstník	Rok	Počet stran	Typ dokumentu	Nezkrácený název dokumentu podle EUR-Lex
32002R0052	L 10 2002	2002	3,2	Nářízení	Nářízení Komise (ES) č. 52/2002 ze dne 11. ledna 2002, kterým se mění nařízení (ES) č. 245/2001, kterým se stanoví prováděcí pravidla k nařízení Rady (ES) č. 1673/2000 o společné organizaci trhu se lnem a konopím pěstovanými na vlákno
32001R1093	L 150 2001	2001	3,6	Nářízení	Nářízení Komise (ES) č. 1093/2001 ze dne 1. června 2001, kterým se mění nařízení (ES) č. 245/2001, kterým se stanoví prováděcí pravidla k nařízení Rady (ES) č. 1673/2000 o společné organizaci trhu se lnem a konopím pěstovanými na vlákno
32005R0873	L 146 2005	2005	4	Nářízení	Nářízení Komise (ES) č. 873/2005 ze dne 9. června 2005, kterým se mění nařízení (ES) č. 245/2001, kterým se stanoví prováděcí pravidla k nařízení Rady (ES) č. 1673/2000 o společné organizaci trhu se lnem a konopím pěstovanými na vlákno
32001R0245	L 35 2001	2001	20,8	Nářízení	Nářízení Komise (ES) č. 245/2001 ze dne 5. února 2001, kterým se stanoví prováděcí pravidla k nařízení Rady (ES) č. 1673/2000 o společné organizaci trhu se lnem a konopím pěstovanými na vlákno
52006DC0125	KOM/2006/0125	2006	22	Zpráva	Zpráva Komise Evropskému parlamentu a Radě o odvětví lnu a konopí
			Celkem:53,6		

### 1.3. Typy dokumentů- druhy právních předpisů



Z tabulek (č. 1 až č. 3) vyplývá, že většina mých textů pochází dle Úředního věstníku z řady L, jedná se tedy o právní předpisy v různých formách. V mém případě se jedná o tyto předpisy: nařízení, směrnice a návrh nařízení. V jednom případě text pochází podle Úředního věstníku z řady C, jde o informace a oznámení, v tomto případě se jedná o stanovisko. Nejčastěji jsou texty vypracovány ve formě nařízení – z 30 právních předpisů je jich 25 nařízením. Pro znázornění jsem vytvořila výsečový graf, který představuje všechny použité formy textů a je opatřen popisky s informacemi o druhu textu, jeho počtu výskytů a jeho procentuálním ekvivalentem.

Graf č. 2: Druhy právních textů



## 2. Zpracování textů, program ParaConc

V této části kapitoly popíši technickou část své práce. Vybrané texty bylo třeba nejdříve „očistit“, aby byly připravené pro nahrání do programu ParaConc a následného vytvoření paralelního korpusu. Dalším důležitým krokem byl samotný výběr klíčových slov a kolokací. Výsledkem této technické části bylo vytvoření databáze vybraných klíčových slov a kolokací v programu MS Access.

### 2.1. Úprava formátu legislativních textů

Vybrané texty jsem si nejdříve zkopírovala ze stránek Eur-Lexu do programu Microsoft Word a následně je uložila podle celexového čísla\_zkratka jazyka (*např. 31971R1696\_CS a 31971R1696\_FR*), ve formátu „.doc“. Poté jsem si texty očistila, to znamená, že jsem odstranila tabulky a různé poznámky. Tyto tabulky, či poznámky se vyskytovaly většinou na koncích textů, někdy i uvnitř textu. Nejčastější příčinou pro odstranění těchto tabulek a poznámek bylo, že buď byly obsaženy pouze v jedné jazykové verzi, nebo byly pro náš výzkum z terminologického hlediska bezvýznamné. Dále jsem texty upravila tak, aby měly fyzicky stejný počet odstavců, tj. stejný počet značek konců odstavců „¶“. Znak konce odstavce by měl v souborech oddělovat odstavce, nadpisy, položky seznamů apod. Následně jsem zkontrolovala přes „Nástroje“ – „Počet slov“, jestli je shodný počet odstavců jak v českých textech, tak i ve francouzských. Pokud se počet odstavců lišil, tak jsem mechanicky pročítala texty a kontrolovala odstavce. Nejčastěji se stávalo, že byl vložen znak odstavce uprostřed věty. Další krok bylo uložení textů ve formátu *prostý text*, čili „.txt“. Při tomto kroku bylo důležité kontrolovat postup ukládání. České texty se ukládaly v kódování Středoevropské jazyky (Windows) a texty francouzské se ukládaly v kódování Západoevropské jazyky (Windows). Takto uložené texty ve formátu *prostý text* byly připraveny pro nahrání do programu ParaConc<sup>2</sup> a vytvoření paralelního korpusu.

### 2.2. Vytvoření paralelního korpusu

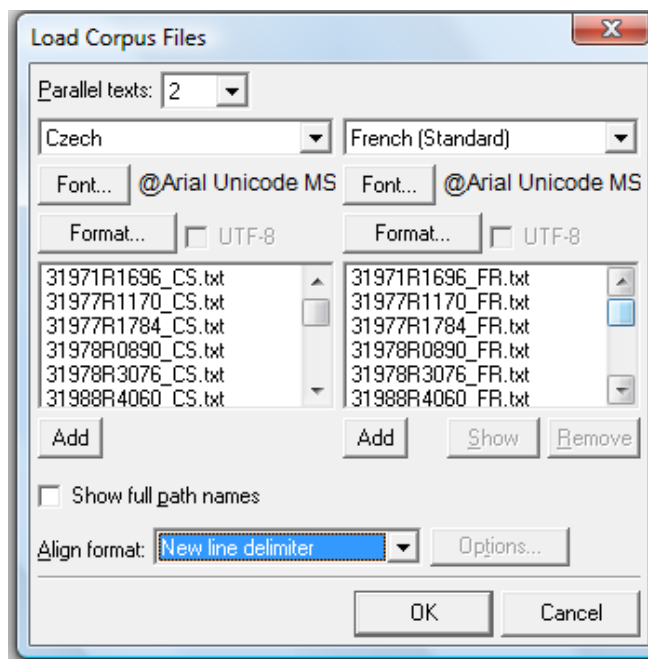
Abychom mohli texty v programu ParaConc otevřít, musíme je nejdříve nahrát. V ParaConcu klikneme na „File“ a vybereme „Load Corpus Files“, kde si vybereme

---

<sup>2</sup> ParaConc je program pro vytváření a prohlížení paralelních korpusů. Dostupné z WWW: [http://utkl.ff.cuni.cz/~rosen/public/IC\\_manual8/node33.html](http://utkl.ff.cuni.cz/~rosen/public/IC_manual8/node33.html) [cit. 17.3.2010].

potřebný počet paralelních textů, pro mou práci dva paralelní texty a následně vybereme potřebné jazyky<sup>3</sup>. ParaConc přijímá až čtyři paralelní texty, které by mohly mít čtyři různé jazyky, nebo jeden původní text a tři různé překlady. Pro vytvoření mého korpusu vybereme vlevo *Czech* a vpravo *French (Standard)* kam postupně nahrajeme „Add“ připravené a uložené texty ve formátu *prostý text*. Abychom ověřili, jestli se texty shodují v počtech odstavců, zaškrtneme pole „Align Format“ a „New Line Delimiter“.

Obr. 1: Nahrání textů v Programu ParaConc



Úspěšné vyhledávání a analýza paralelních textů závisí na souladu textových segmentů v každém jazyce korpusu. Zarovnání (*Alignment*) je údaj o ekvivalentnosti textových segmentů ve dvou jazycích. ParaConc jako základní segment zarovnání používá větu, avšak já jsem pro svůj korpus použila zarovnání na odstavce.

Odlišný počet odstavců nám ParaConc ohlásí. V tom případě se musíme vrátit o krok zpět k uloženým textům ve formátu „.txt“ a zkontrolovat je a případné nedostatky odstranit. Poté texty opět nahrajeme do programu a zkontrolujeme, jestli jsou odstavce shodné.

<sup>3</sup> Pokud příslušný jazyk mezi nabízenými možnostmi nenajdete, je třeba příslušné národní prostředí do systému doinstalovat. Dostupné z WWW: [http://utkl.ff.cuni.cz/~rosen/public/IC\\_manual8/node16.html](http://utkl.ff.cuni.cz/~rosen/public/IC_manual8/node16.html) [cit. 17.3.2010].

Může se nám také stát, že nám nahrané texty zobrazují abnormální znaky. V tom případě zkontrolujeme nastavení kódování. V menu Load Corpus files klikneme na „Font“ a zkontrolujeme nastavení pro české texty, tj. Středoevropský a pro francouzské texty Západní. Pokud ani toto nepomůže, musíme se vrátit do textů uložených ve formátu Word a zkontrolovat písmo. Mohlo se nám stát, že jsme zvolili druh písma, který neobsahoval všechny znakové sady. Vhodné je tedy volit druh písma Arial.

Jakmile se soubory nahrají, na obrazovce menu se objeví některé nové položky týkající se analýzy a zobrazení textu. Jedná se o tyto funkce: soubor „FILE“, hledat „SEARCH“, četnost „FREQUENCY“ a informace o programu „INFO“. Navíc můžeme získat informace v levém dolním rohu okna o počtu nahraných (přidaných) souborů a v pravém dolním rohu počet slov pro dva texty, které jsme do ParaConcu nahráli.

### 2.3. Výběr klíčových slov a kolokací

Z těchto textů jsem si dále vybrala 32 klíčových slov. Klíčová slova jsem vybírala nejprve podle frekvence výskytů (100 - 400 výskytů), poté jsem si vybírala i méně frekventovaná slova (např. slova, která se v textech objevují méně než desetkrát). K těmto klíčovým slovům jsem následně dohledala 125 kolokací, které jsou vhodné pro mou práci. Celkem tedy budu pracovat se 157mi termíny a kolokacemi.

Základní vyhledávání je poměrně jednoduché: lze zadat slovo nebo fráze, včetně jednoduchých zástupných znaků v případě potřeby. Symboly působící jako zástupné znaky jsou uživateli definovány. Pro vyhledávání klíčových slov nebo kolokací lze využít těchto znaků:

- ? – nahrazuje jeden znak
- % – žádný nebo jeden znak
- \* – nahrazující více znaků

Já jsem např. pro slovo industrie vyhledávala tak, že jsem si do *Search* (French Standard) zadala \*industrie, kdy mi ParaConc našel toto slovo ve 39 výrazech. Kdy mezi nejčastěji se vyskytujícími převládá výraz se členem, čili l'industrie. Pro výraz emballage jsem vyhledávala stejným způsobem a ParaConc našel celkem 34 výrazů,

zahrnujíc také podstatné jméno s určitým členem l'emballage a podstatné jméno s předložkou *de*, čili d'emballage.

Pro vyhledávání klíčových slov jsem nejprve použila v programu ParaConc funkce „*Frequency*“ – „*Frequency Order*“ – „*French (Standard)*“, z kterých jsem se pokusila vybrat 20 nejčastějších klíčových slov. Důležité bylo, aby vybraná klíčová slova byla vhodná pro mou oblast výzkumu, proto jsem se vyhýbala termínům, které byly sice nejfrekventovanější, avšak pro mou práci nepříliš vhodné (jednalo se o slova jako např. *règlement*, *article* apod.), kdy např. termín *article* označuje jednotlivé články ve smlouvě a proto netvoří kolokace. Termín *règlement* označuje, o jaký typ dokumentu se jedná. V této základní části program ParaConc vyhledává ze sto nejfrekventovanějších slov. Ve funkci „*Frequency and Collocation Options*“ je možnost nastavení vyhledávání např. z pěti tisíc nejčastějších slov.

Vybraná slova jsem si přepsala do programu Microsoft Access, který je přijatelný zejména díky tomu, že dokáže zpracovat velké množství dat a zároveň díky propojitelnosti mezi tabulkami, které jsem si v programu vytvořila.

Seznam vybraných klíčových slov najdeme v tabulce níže (*Tabulka č. 4: Seznam vybraných klíčových slov*), která obsahuje také slovní druh, RANK, což je pořadí ve frekvenčním seznamu, které zjistíme použitím funkce Hot words a nakonec je v tabulce obsažen celkový počet výskytů (i procentuelně). Tato klíčová slova jsem seřadila podle abecedy. Vybraná klíčová slova jsou všechna podstatná jména, kde oba rody<sup>4</sup> podstatných jmen jsou zastoupeny téměř stejným dílem. Podstatná jména ženského rodu reprezentují 50% (tj. 16 slov) a podstatná jména mužského rodu představují zbylou část, tedy 50% (tj. 16 slov). V obou případech jsem brala v potaz i podstatná jména v množném čísle.

*Tabulka č. 4: Seznam vybraných klíčových slov*

Termíny			
Termín	Slovní druh	RANK	Počet výskytů
accise	s.f.	5,87	5 (0,0055%)
certification	s.f.	17,76	77 (0,0852%)
comité	s.m.	37,21	45 (0,0498%)
commerce	s.m.	19,73	22 (0,0243%)

<sup>4</sup> na rozdíl od českého jazyka, který má tři rody (mužský, ženský a střední) má francouzský jazyk pouze rody dva, a to mužský a ženský.

Termíny			
Termín	Slovní druh	RANK	Počet výskytů
commercialisation	s.f.	32,10	135 (0,1493%)
cônes	s.m.pl.	88,79	36 (0,0398%)
contrôle	s.m.	11,53	192 (0,2124%)
emballage	s.m.	18,49	34 (0,0376%)
feuilles	s.m.pl.	70,48	51 (0,0564%)
fibres	s.f.pl.	79,16	326 (0,3606%)
graines	s.f.pl.	20,32	33 (0,0365%)
groupement	s.m.	184,83	202 (0,2235%)
hausse	s.f.	9,94	18 (0,0199%)
houblon	s.m.	91,31	388 (0,4182%)
humidité	s.f.	10,68	35 (0,0387%)
champ	s.m.	7,85	13 (0,0144%)
chanvre	s.m.	71,25	186 (0,2058%)
industrie	s.f.	12,58	39 (0,0431%)
lin	s.m.	12,23	229 (0,2533%)
mouture	s.f.	36,97	5 (0,0055%)
politique	s.f.	62,91	44 (0,0487%)
poudre	s.m.	48,53	28 (0,0310%)
prime	s.f.	158,13	129 (0,1427%)
prix	s.m.	24,46	158 (0,1748%)
producteur	s.m.	62,25	131 (0,1449%)
production	s.f.	22,21	399 (0,4414%)
quotas	s.m.pl.	81,99	112 (0,1239%)
récolte	s.f.	51,93	167 (0,1847%)
superficie	s.f.	41,63	40 (0,0443%)
tabac	s.m.	33,33	331 (0,3662%)
transformation	s.f.	79,21	299 (0,3308%)
variétés	s.f.pl.	15,95	122 (0,1350%)

K vybraným termínům jsem si postupně dohledávala vhodná slovní spojení neboli kolokace.

### 2.3.1. Definice pojmů

- **Termínem** chápeme lexikální jednotku, která tvoří odborné pojmenování v soustavě daného oboru. (Akademický slovník cizích slov, 1995:756)

- **Kolokace** je dle F. Čermáka (2004:254) definována takto: „*Smysluplné spojení lexémů/slov, lexikální syntagma, zvl. v podobě víceslovného pojmenování, jehož vznik je podmíněn jejich vzájemnou kolokabilitou<sup>5</sup>, a tedy i kompatibilitou<sup>6</sup>.*“
- Za **klíčové slovo** ve své práci považuji lexikální jednotku, která je úzce spjata s tématem „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“ a je součástí kolokace.

### 2.3.1.1. Pojem kolokace

Termín *kolokace* má původ v latinském slově *collocare*. Za otce kolokací bývá obvykle považován britský lingvista J. R. Firth, který termín kolokace zavedl do jazykovědy již ve 30. letech 20. století. Do češtiny byl tento pojem pravděpodobně zaveden Františkem Čermákem., který se kolokacemi zabývá. Kolokace obecně odkazuje na vyjádření slov, která jsou často používána společně např. „bílá káva“. Termín kolokace nahradil dnes již poměrně málo používané *slovní spojení*.

V korpusech můžeme kolokace vyhledávat dvěma způsoby a to tak, že jsou buď základní složkou korpusového manažeru např. české Bonito, nebo pomocí řady speciálních programů např. Collocate, WordSmith aj. (srov. Čermák, 2006:12)

#### 2.3.1.1.1. Typy kolokací

F. Čermák vytvořil klasifikaci typů lexikálních kombinací.<sup>7</sup> Při sestavování této klasifikace se F. Čermák řídil následujícími hledisky:

- (A) systém text (langue-parole a tedy ustálený-neustálený)
- (B) pravidelný-nepravidelný (formálně i sémanticky)
- (C) kolokace-nekolokace (jiné syntagma)
- (D) proprium-neproprium

<sup>5</sup> kombinatorická schopnost, resp. syntagmatický potenciál slova (ČERMÁK, František. *Studie z korpusové lingvistiky : Kolokace: Lidové noviny*, 2006:11. Kolokace v lingvistice. ISBN 80-7106-863-2.)

<sup>6</sup> sémantická slučitelnost (viz. výše)

<sup>7</sup> ČERMÁK, František. *Studie z korpusové lingvistiky : Kolokace : Lidové noviny*, 2006:12-13. Kolokace v lingvistice. ISBN 80-7106-863-2.

Na základě těchto pravidel F. Čermák rozděluje kolokace na *systemové*, *textové* a *textové-systemové*.

### **A Systemové**

**1 pravidelné**                    **a: termínové kolokace** (víceslovné termíny)

**b: propriální kolokace** (víceslovná propria)

**2 nepravidelné:**            **idiomatické kolokace** (idiomy a frazémy)

#### **extenze a přechody**

### **B textové**

**3 pravidelné**                    **a: běžné kolokace (gram-sém. kombinace)**

**b: analytické kombinace tvarů** (analytické formy)

**4 nepravidelné**                **a: individuální metaforické kolokace** (autorské metafor)

**b: náhodné kombinace sousední**

**c: jiné kombinace** (blábol)

### **C textové-systemové**

**5**                                    **běžné kolokace uzuální**

Případně můžeme brát v potaz i další přechodné typy, zejména pak kombinace mezi A1a a B3a.



Pro příklady a vysvětlení jednotlivých kombinací budu vycházet z článku F. Čermáka - Syntagmatika slovníku: typy lexikálních kombinací, který byl publikován ve sborníku *Čeština – univerzália a specifika 3*.<sup>8</sup>

### ***Termínové kolokace***

Pro termínové kolokace uvádí F. Čermák jako příklady: „*cestovní kancelář*“, „*kyselina sírová*“, „*lodní doprava*“, „*to co*“ a „*až když*“.<sup>9</sup>

Zde F. Čermák upozorňuje na problém určení, jestli se stále ještě/už jedná o *termín* nebo je to ještě/už *ustálenost*. Nabízí zde však řešení, kterého dosáhneme aplikací *prototypového přístupu*, kdy rozlišíme centrum a periferii. Jako příklad prototypového příkladu uvádí např. „*bílý chléb*“, který je termínem pro pekaře (lidí z daného oboru), ale ne pro neutrální uživatele.

Pro příklady kolokací typů „*to co*“, „*až když*“ F. Čermák podotýká, že jejich počet je ohraničen, v češtině se jich vyskytuje jen několik stovek.

### ***Propriální kolokace***

Propriální kolokace jsou málo prozkoumané, týkají se v první řadě problematiky antroponym, kde také narážejí na problém ustálenosti. F. Čermák jako příklady uvádí: „*Kanárské ostrovy*“, „*Středozemní moře*“, „*Velká Británie*“. Dále F. Čermák zmiňuje, že proprium bude vždy alespoň pro malý okruh lidí ustálený.

### ***Idiomatické kolokace***

Jako příklady idiomatických kolokací F. Čermák uvádí „*bez nadsázky*“, „*ležet ladem*“, „*jen aby*“, „*stará dobrá Anglie*“. Tyto kolokace jsou založeny na sémantických formálních anomáliích, nepravidelnostech a řeší stejně jako kolokace termínové otázku *ustálenosti*. F. Čermák doporučuje prototypové řešení a v případě nejasností/sporů případy ověřit přímo na korpusu.

---

<sup>8</sup> ČERMÁK, František. Syntagmatika slovníku: typy lexikálních kombinací. In HLADKÁ, Z.; KARLÍK, P. *Čeština - univerzália a specifika 3*. 2001:223-232, ISBN 80-210-2532-8. Dostupné z WWW: <<https://mathesius.ff.cuni.cz/ucnk/doc/synkol.rtf>>. [cit. 15.4.10]

<sup>9</sup> Dostupné z WWW: <https://mathesius.ff.cuni.cz/ucnk/doc/synkol.rtf> [cit. 15.4.2010]

Dále F. Čermák do této skupiny řadí i *extenze a přechody*. Vedle extenze do proprií např. „stará dobrá Anglie“ zmiňuje F. Čermák i extenzi do termínů např. „černá díra“ – v tomto případě extenze do astronomie.

F. Čermák ještě připomíná, že do této skupiny patří jako periferní podtyp kvazifrazému velmi početná skupina *verbonominálních kvazifrazémů* např. „věnovat pozornost“.

### ***Běžné kolokace***

Oblast těchto kolokací je velmi bohatá a rozmanitá, přičemž patří do našeho všedního vyjadřování. Jako příklady F. Čermák uvádí „letní dovolená“, „snadná odpověď“, „teplé podnebí“, „starověké památky“, „místní víno“, „jižní pobřeží“, „největší ostrov“, „právě ten“. Vyskytují se zde případy specifikace, kdy jeden lexém upřesňuje druhý (např. „letní dovolená“, kde „letní“ upřesňuje lexém „dovolená“ z této stránky se jedná o tzv. **spatiotemporální** (místně-časové) určení.

### ***Analytické kombinace tvarů***

V tomto případě F. Čermák uvádí, že se nejedná o kolokace, nýbrž o pravidelné, ale neustálené formy jednoslovných lexémů, např. „šel by“, „byl zapsán“, „vzpomínající si“.

### ***Individuální metaforické kolokace***

K tomuto typu kolokací F. Čermák podotýká, že by neměly být primárně spojovány s básnickými texty, protože těch je ve skutečnosti málo. Dále podotýká, že je tato oblast neprozkoumána, byť je velmi rozsáhlá a občas užita prakticky každým mluvčím. Jako hlavní problém F. Čermák vidí jejich odlišení od *běžných kolokací*. Jako příklady uvádí např. „třeskutě vtipný“, „virové hrátky“.

### ***Náhodné kombinace sousední***

V tomto případě se jedná o čistě náhodné a často i nesmyslné kombinace slov, které se vedle sebe ocitly v důsledku slovosledu a tím pádem je nemůžeme chápat ani jako kolokace ani jako víceslovné lexémy. Např. „(vývody) vzduchotechniky uvnitř (bytu)“, „že v“, „kterou jistě“.

### ***Jiné kombinace***

Tuto skupinu sem F. Čermák zařazuje jen spíše z důvodu úplnosti, protože se jedná o kombinace, které nespádají do žádného typu kolokací zmíněných výše.

### ***Běžné kolokace uzuální***

Jedná se o kolokace, které mají pohyblivý charakter a mohou být různými lidmi chápány různě. Tyto kolokace jsou na hranici mezi ***termínovými kolokacemi*** a ***běžnými kolokacemi***. Příklady, které F. Čermák pro tuto skupinu uvádí, jsou: „*prát prádlo*“, „*sušit prádlo*“, „*žehlit prádlo*“, „*nakrájet nadrobno*“, „*umýt si ruce*“, „*nastoupit do tramvaje*“.

Po prostudování kolokací byla naprostá většina kolokací zhodnocena jako termínové kolokace (např. „*tabac brut*“, „*paille de chanvre*“ a další). Dále byly kolokace zhodnoceny jako běžné kolokace (např. „*prix de vente*“, „*commerce des produits*“, „*industrie textile*“ a další). Ojediněle se jedná také o náhodné kombinace sousední (např. „*règles relatives à la commercialisation*“, „*situation concurrentielle sur le marché du tabac*“, „*produit obtenu par mouture du houblon*“), tím pádem se ne vždy bude jednat o skutečné kolokace, ale o více samostatných lexikálních jednotek stojících vedle sebe.

Rozdělení kolokací dle morfologických struktur se budu věnovat v kapitole 4.1, v 5. kapitole pak zhodnotím české překlady jednotlivých kolokací.

## 3. Korpus, korpusová lingvistika

### 3.1. Korpusová lingvistika

Korpusová lingvistika je poměrně mladá disciplína. Ve své novější, počítačové podobě se korpusová lingvistika rozšířila začátkem 90. let minulého století. V České republice se ale tento pojem objevuje, až od roku 1994, kdy byl na filozofické fakultě Univerzity Karlovy založen Ústav Českého národního korpusu, který vede profesor František Čermák.

„Korpusová lingvistika je taková část lingvistiky, která systematicky pracuje s korpusem a jeho nástroji, resp. studuje zásady a praxi práce s ním s cílem lepšího poznání funkce a struktury jazyka, jaké až dosud nebylo možné.“ (Čermák, 2000:17)

### 3.2. Korpus

Pojem korpus patří mezi starší termíny, kdy se již dříve tento pojem vyskytoval v různých oblastech a významech. V Akademickém slovníku cizích slov je korpus definován jako *soubor jazykových (analyzovaných a vykládaných) materiálů (vět, textů apod.)*. (Akademický slovník cizích slov I. (A-K), 1995:425). Avšak v korpusové lingvistice se korpusem podle Čermáka chápe *zejména takový ucelený soubor textů, který je sestavený s přihlédnutím ke svému cíli reprezentativním způsobem, je obhospodařovaný počítačově a zpracováván souborem korpusových metod*. (Čermák, 2006:9)

*Korpus je absolutně jak nejbohatší a nejspolehlivější, tak i nejrealističtější a nejautentičtější zdroj jazykových dat, a tedy i taková možnost studia a poznání jazyka (a světa za ním)*. (Čermák, 2004:100).

Obrovská rychlost, dostupnost a pohodlné používání výpočetní techniky pomohly k rozvoji korpusu. V minulosti se jazykový materiál sbíral ručně, což vyžadovalo velké množství pracovníků a času. Dříve se takto sbíral např. materiál pro výkladové slovníky.

### 3.3. Typy korpusů

Základní dělení korpusů je lingvistické, kdy korpusy dělíme jednak na psané a mluvené a dále pak na synchronní a diachronní. Avšak existuje i další dělení korpusů např. na paralelní, nářeční atd.

### 3.3.1. Textový korpus

Pro pojem jazyková databanka se již tradičně ve světě užívá označení textový korpus.<sup>10</sup> Textové korpusy byly již vytvořeny pro řadu evropských jazyků. Největší z nich je britský, nazývaný The Bank of English, který obsahuje 350 miliónů slovních tvarů a zahrnuje jak psaný, tak mluvený jazyk.

V České republice je textový korpus zastoupen Českým národním korpusem<sup>11</sup>, jenž vznikl v roce 1991. Již v té době čítal 100 miliónů slovních tvarů a jeho část, která má přes 20 miliónů slov, je již od roku 1996 veřejně přístupná na Internetu (na adrese <http://ucnk.ff.cuni.cz/>).

### 3.3.2. Paralelní korpus

Paralelní korpusy jsou takové korpusy, které slouží zejména lexikografům, překladatelům, literárním vědcům, učitelům nebo studentům cizích jazyků.

F. Čermák definuje paralelní korpus takto: „*Paralelní korpusy jsou korpusy dvou nebo více jazyků vytvářené z překladů, obsahují tedy vždy jednak texty originální, jednak jejich jinojazyčné mutace.*“<sup>12</sup>

Paralelní korpus se v poslední době stává čím dál více využívaným korpusem a to zejména pro oblast překladatelství a lexikografie. Zejména po vstupu České republiky do Evropské unie získávají paralelní korpusy na popularitě, protože slouží k porovnání překladů legislativních nebo jiných textů.

Pro svou práci využijí právě tohoto korpusu, jedná se o paralelní korpus legislativních textů ve francouzském a českém jazyce.

---

<sup>10</sup> Blatná R.; *Jazyková databanka neboli korpus*, In: Vesmír, 1997, č. 12 str. 670-671 Dostupné z WWW: [www.vesmir.cz/files/file/199/aid/3118](http://www.vesmir.cz/files/file/199/aid/3118) [cit. 27.4.10]

<sup>11</sup> Český národní korpus. Dostupné z WWW: <http://ucnk.ff.cuni.cz> [cit. 22.3.10]

<sup>12</sup> ČERMÁK, František; SCHMIEDTOVÁ, Věra. Český národní korpus - základní charakteristika a širší souvislosti. In *Národní knihovna : knihovnická revue*. č3., 2004:154 [cit. 27.4.10]. Dostupné z WWW: <[knihovna.nkp.cz/pdf/0403/0403152.pdf](http://knihovna.nkp.cz/pdf/0403/0403152.pdf)>.

Můžeme rozlišit dva základní typy paralelních korpusů:

- srovnatelný
- překladový

Termín srovnatelný korpus odkazuje na texty ve dvou jazycích, které jsou si podobné v obsahu, ale nejsou překlady.

### 3.3.3. *Synchronní korpus*

Synchronní korpus je psaný korpus, který je založen na současných textech. Analýza těchto textů je nejpotřebnější a zároveň elektronicky nejdostupnější.

Rozsahem tento korpus zahrnuje několik posledních desetiletí, kdy je brán ohled na aktuálnost a životnost textů, nebo na významné vnější (společenské) přeměny.

### 3.3.4. *Diachronní korpus*

Diachronní korpus oproti korpusu synchronnímu pokrývá několik vývojových stádií jazyka, nebo celý jeho vývoj. Zásadním rozdílem mezi synchronními korpusy a diachronními korpusy je rozdílné pojetí reprezentativnosti. V případě diachronních korpusů lze brát v potaz pouze starší dochované texty, jejichž výběr žánrů bývá často omezen a nevyvážen. V těchto případech převažují texty náboženské, legendické, veršované a jinak neobvyklé.

### 3.3.5. *Mluvený korpus*

Mluvené korpusy se řadí mezi novější korpusy. Vzhledem k obtížnosti mluvených korpusů (časové i finanční), která je mnohem větší než u psaných korpusů, je jich tím pádem také méně a jsou i menší. V Evropě je v dnešní době obtížné najít jazyk, který by neměl psaný korpus, přičemž některé korpusy jsou i dost velké (např. angličtina, němčina, čeština, španělština aj.), situace v oblasti mluvených korpusů je prakticky obrácená. Mluveným korpusem disponuje jen velmi málo jazyků a jejich

rozsah i využitelnost je omezená. Největší z mluvených korpusů nepřesahují 9-10 milionů slov.<sup>13</sup>

### Typy mluvených korpusů<sup>14</sup>

František Čermák rozděluje mluvené korpusy na dva typy a to na specializovaný korpus a obecný korpus.

*Specializovaný korpus* – bývá nějakým způsobem omezený. Může být omezený sociálním nebo jazykovým jevem, na který se zaměřuje, popř. tématicky, žánrově apod.

*Obecný korpus* – má být reprezentativní pro žánry a typy mluveného jazyka stejným způsobem jako je pro celý jazyk reprezentativní korpus psaný. Je co do povahy v různé míře závislý na jazyku, pro který je vytvářený. Podoba je předurčena typologicky a sociologicky.

#### 3.3.5.1. *Cizí mluvené korpusy*

Mezi největší mluvený korpus patří *British National Corpus (BNC)*. Nahrávky do tohoto korpusu vznikly v rozmezí let 1991 – 1995 a pocházejí ze dvou zdrojů:

- výběrem ze známějších žánrů, kam patřily např. přednášky, kurzy a škola, zprávy, obchodní jednání, interview, politické projevy a další.
- vzorkem žánrově a obsahově nspecifickým, tj. 2000 hodin nahrávek od více než 100 dobrovolníků z celé země, které byly pořízeny v 38 oblastech muži i ženami a věku v rozmezí 15 – 60 let, kdy dobrovolník nahrával celý svůj den, tj. všechny hovory, které tento den měl.

Novějším korpusem, dokončeným v roce 2004, a téměř stejně rozsáhlým je *Korpus mluvené nizozemštiny (The Spoken Dutch Corpus)*.

Mluvených korpusů v angličtině existuje sice víc, ale jsou i poměrně menší. Jedním z těchto menších korpusů je nejstarší anglický korpus z roku 1959 *London-Lund*

---

<sup>13</sup> ČERMÁK, František. *Korpusová lingvistika: Stav a modelové přístupy*: Lidové noviny, 2006:54. Mluvené korpusy, ISBN 80-7106-865-9.

<sup>14</sup> ČERMÁK, František. *Korpusová lingvistika: Stav a modelové přístupy*: Lidové noviny, 2006:53-54. Mluvené korpusy, ISBN 80-7106-865-9.

*Corpus*, který má rozsah 0,5 milionu slov. R. Quirk tento korpus vybudoval na 87 textech ze *Survey of English Usage*, které později byly doplněny J. Svartvikem o 13 dalších textů. Dalšími anglickými korpusy jsou např. *BASE*, *British Academic Spoken English corpus*, *American National Corpus* nebo *The Longman Spoken American Corpus* a další.

### 3.3.5.2. *České mluvené korpusy*

V letech 1988 – 1995 vznikl pro češtinu *Pražský mluvený korpus (PMK)*, který je součástí Českého národního korpusu a má rozsah 819 000 slov (včetně diakritiky), ale roste dál. Obsahově patří Pražský mluvený korpus mezi obecné korpusy, neboť je založený na nahrávkách volných nepřipravených dialogů a odpovědí 304 mluvčích v Praze a okolí.

Na Pražský mluvený korpus navazuje *Brněnský mluvený korpus*, který vznikl v letech 1994 – 1999 stejnou metodou. Obsahuje 0,6 milionu slov.

### 3.4. *Využití korpusů*

Využití korpusů je dnes rozšířeno do mnoha oblastí a oborů. Ačkoli je využití korpusů mezioborové, tak mezi disciplíny, které korpusy využívají nejčastěji, patří právě lexikografie a překladatelství.

Korpus, kterým se budu zabývat v této práci, bude, jak již bylo zmíněno paralelní. Z hlediska reprezentativnosti korpusu patří texty mezi reprezentativní, neboť se jedná o legislativní texty Evropské unie, jako jsou nařízení, stanoviska apod., z hlediska velikosti se jedná o malý korpus, neboť obsahuje pouze několik desítek tisíc slov.



## 4. Práce s texty

V první části této kapitoly rozřadím kolokace z hlediska morfologických struktur. V další kapitole (5. kapitola) se budu zabývat správností, či nesprávností překladů všech kolokací k vybraným termínům.

### 4.1. Morfologické struktury kolokací

Pro rozdělení kolokací budu vycházet z práce Michela Mathieu-Colase (1996), který francouzská kompozita rozděluje na 17 základních tříd a 8 doplňkových struktur tříd, celkem tedy na 25 tříd. Tyto třídy jsou dále ještě děleny na mnoho dalších podtříd, celkem tedy francouzská kompozita čítají až 700 morfologických struktur (Mathieu-Colas; 1996:71).

### 4.2. Struktury kolokací pro vybrané texty z oblasti: „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“

Pro zachování přehlednosti v dělení kolokací budu vycházet ze zkratk, které se nacházejí v příloze (viz. *Seznam zkratk*).

*Tabulka č. 5.1: Struktura kolokací N+Adj. (10. třída)*

Do této struktury kolokací jsem zařadila také kolokaci „*prix plus élevés*“, kdy „*plus élevés*“ budu považovat za jedno slovo (v tomto případě adjektivum) a ne za spojení příslovce + přídavné jméno.

Kolokace	Výskytů	Struktura kolokace
commerce intérieur	1	N+Adj.
politique sociale	1	N+Adj.
accise minimale	2	N+Adj.
politique sanitaire	2	N+Adj.
politique fiscale	2	N+Adj.
emballage final	2	N+Adj.
prix raisonnables	2	N+Adj.
tabac naturel	2	N+Adj.
producteur individuel	24	N+Adj.
superficie cultivée	3	N+Adj.

Kolokace	Výskytů	Struktura kolokace
lin textile	3	N+Adj.
politique budgétaire	4	N+Adj.
prix plus élevé	4	N+Adj.
industrie textile	5	N+Adj.
tabac emballé	6	N+Adj.
tabac brut	61	N+Adj.
commerce illicite	7	N+Adj.

Tabulka č. 5.2: Struktura kolokací N+de+X (12. třída)

Do této třídy kolokací spadají především kolokace se strukturou N+de+N, tudíž se vyskytují přesně dvě podstatná jména.

Kolokace	Výskytů	Struktura kolokace
industrie du tabac	1	N+de+N
producteur de houblon	1	N+de+N
répartition de quotas	1	N+de+N
écoulement de la récolte	1	N+de+Det.+N
fluctuations des prix	1	N+de+N
conservation du tabac	1	N+de+N
rouleaux de tabac	1	N+de+N
déchets de tabac	1	N+de+N
fixation des prix	1	N+de+N
prix d'objectif	1	N+de+N
enregistrement des superficies	1	N+de+N
politique de sécurité	1	N+de+N
stade de commercialisation	1	N+de+N
évolution des prix	1	N+de+N
graines de chanvre	10	N+de+N
variétés de tabac	10	N+de+N
contrat de transformation	10	N+de+N
année de récolte	11	N+de+N
prix de vente	11	N+de+N
livraison du tabac	12	N+de+N
prix des cigarettes	12	N+de+N
montant de la prime	14	N+de+Det.+N
poudre de houblon	15	N+de+N
organisme de contrôle	15	N+de+N
procédure de certification	16	N+de+N
attestation de contrôle	17	N+de+N
pailles de lin	17	N+de+N
quantités de fibres	18	N+de+N
organisme de certification	2	N+de+N
centres de certification	2	N+de+N
livraison de houblon	2	N+de+N

Kolokace	Výskytů	Struktura kolokace
hausse du prix	2	N+de+N
production de paille	2	N+de+N
superficie de lin	2	N+de+N
prix de marché	2	N+de+N
écoulement du tabac	2	N+de+N
feuilles de cône	2	N+de+N
industrie des fibres	2	N+de+N
poids du tabac	2	N+de+N
prix des produits	2	N+de+N
taux de humidité	21	N+de+N
prix d'achat	23	N+de+N
secteur du lin	23	N+de+N
secteur du houblon	28	N+de+N
unité d'emballage	3	N+de+N
régions de production	3	N+de+N
transformation du tabac	3	N+de+N
récolte des pailles	3	N+de+N
distribution de quotas	3	N+de+N
augmentation des prix	3	N+de+N
quotas de production	33	N+de+N
industrie du papier	4	N+de+N
quotas de transformation	4	N+de+N
politique de qualité	4	N+de+N
paille de chanvre	4	N+de+N
commercialisation du houblon	4	N+de+N
fibres de chanvre	49	N+de+N
champ d'activité	5	N+de+N
commerce des produits	5	N+de+N
mouture de houblon	5	N+de+N
régime de certification	5	N+de+N
production de fibres	50	N+de+N
groupe de variétés	53	N+de+N
politique de santé	6	N+de+N
transformation du houblon	6	N+de+N
campagne de commercialisation	62	N+de+N
groupement de producteurs	7	N+de+N
régime de prime	7	N+de+N
régime de quotas	8	N+de+N
rachat de quotas	9	N+de+N

Tabulka č. 5.3: Struktura kolokací N+à+X (13. třída)

Kolokace	Výskytů	Struktura kolokace
droit à la prime	1	N+à+Det.+N
prime au producteur	3	N+à+Det.+N

Kolokace	Výskytů	Struktura kolokace
chasse à la prime	3	N+à+Det.+N

Tabulka č. 5.4: Struktura kolokací N+en+X (14. třída)

Kolokace	Výskytů	Struktura kolokace
tabac en feuilles	20	N+en+N
houblon en cônes	26	N+en+N

Tabulka č. 5.5: Struktura kolokací N+autre prép.+X (15. třída)

Kolokace	Výskytů	Struktura kolokace
houblon avec graines	1	N+avec+N
lutte contre le tabagisme	1	N+contre+Det.+N

Tabulka č. 5.6: Struktura kolokací N+Adj.+X (19. třída)

Po zhodnocení do této třídy řadím i kolokaci „du lin et du chanvre destinés à la production de fibres“, kdy „du lin et du chanvre“ budu brát za nominální syntagma.

Kolokace	Výskytů	Struktura kolokace
Organisation Mondiale du Commerce	1	N+Adj.+de+N
règles relatives à la commercialisation	1	N+Adj.+à+Det.+N
situation concurrentielle sur le marché du tabac	1	N+Adj.+sur+Det.+SN
exigences minimales de commercialisation	11	N+Adj.+de+N
Comité économique et social	13	N+Adj.+et+Adj.
partie variable de la prime	17	N+Adj.+de+Det.+N
superficies plantées en houblon	2	N+Adj.+en+N
règles générales de la certification	2	N+Adj.+de+Det.+N
politique agricole commune	22	N+Adj.+Adj.
réserve nationale de quotas	3	N+Adj.+de+N
État membre producteur	4	N+Adj.+Adj.
limites minimales de commercialisation	4	N+Adj.+de+N
prix moyen pondéré	4	N+Adj.+Adj.
produit obtenu par mouture du houblon	5	N+Adj.+par+SN
extraits végétaux de houblon	6	N+Adj.+de+N
Fonds communautaire du tabac	6	N+Adj.+de+N
fibres courtes de lin	75	N+Adj.+de+N
du lin et du chanvre destinés à la production de fibres	19	SN+Adj.+à+Det.+SN

Tabulka č. 5.7: Struktura kolokací N+de+N+X (21. třída)

Do této třídy řadím kolokace, ve kterých se objevují více než 2 podstatná doplněná buď dalším podstatným, nebo přídavným jménem.

Kolokace	Výskytů	Struktura kolokace
commerce des produits agricoles	1	N+de+N+Adj.
groupement de producteurs de houblon	1	N+de+SN
comité de gestion des fibres naturelles	2	N+de+SN
comité de gestion du lin et du chanvre	2	N+de+SN
production de fibres de chanvre	2	N+de+SN
importations de graines de chanvre	2	N+de+SN
règles de commercialisation arrêtées pour le houblon	2	N+de+SN
plante du houblon grimpant	2	N+de+N+Adj.
hausse des droits d'accises	3	N+de+SN
attestations de quotas de production	4	N+de+SN
numéro de référence de certification	4	N+de+SN
comité de gestion du tabac	5	N+de+SN
comité de gestion du houblon	8	N+de+SN

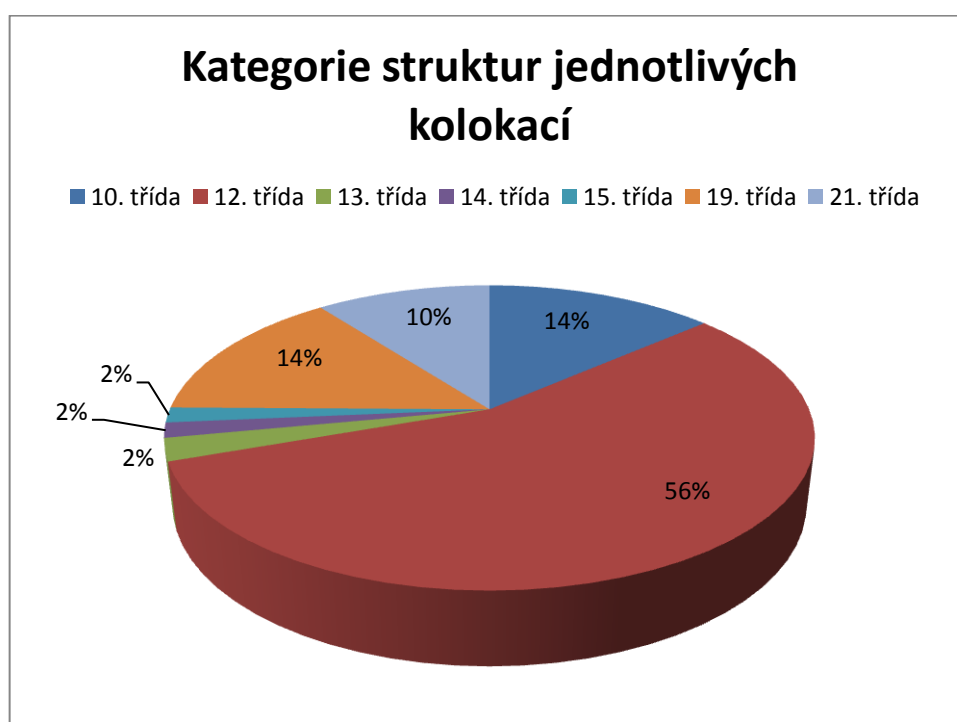
#### 4.3. Závěr

Na závěr bych ráda shrnula výsledky struktur kolokací. Předmětem zkoumání bylo všech 125 kolokací. Tyto kolokace jsem rozčlenila do 7 tříd dělení struktur kolokací ze všech 25 tříd, které nabízí ve své práci Mathieu-Colase. Zjistila jsem, že nejvíce kolokací z mé databáze je z 12. třídy, tj. ze struktury „N+de+X“ a je zastoupena celkem 70 kolokacemi. Procentuelně vyjádřeno je tato kategorie zastoupena u 56% mých kolokací. Druhá a třetí nejpočetnější struktura jsou téměř vyrovnány, avšak 19. třída, čili struktura kolokací „N+Adj.+X“, čítá celkem 18 kolokací, což představuje přibližně 14% a třetí nejpočetnější skupina, 10. třída, struktura kolokací „N+Adj.“ má 17 kolokací, přibližně 14%. Čtvrtou nejpočetnější skupinou je 21. třída, tedy struktura kolokací „N+de+N“, která má 13 kolokací představujících 10%. Zbylé tři skupiny jsou téměř vyrovnány. Jedná se o skupiny, které mají méně než pět kolokací. Tyto skupiny obsahují jednu po třech kolokacích 13. třídy, struktura kolokací „N+à+X“, představující 2% a dvakrát po dvou kolokacích, 14. třída, struktura kolokací „N+en+N“ a 15. třídy „N+autre prép.+X“, které také představují 2%. Pro větší přehlednost jsem vytvořila tabulku i graf. Tabulka zobrazuje v jednom sloupečku struktury kolokací a v druhém sloupečku počet kolokací pro danou strukturu. Graf zobrazuje procentuelní zastoupení daných tříd z celkového počtu 125 kolokací.

Tabulka č. 5.9: Přehled struktur kolokací

Struktura	Počet
N+de+X	70
N+Adj.+X	18
N+Adj.	17
N+de+N	13
N+à+X	3
N+en+X	2
N+autre prép.+X	2

Graf č. 3: Kategorie struktur kolokací



## 5. Překlady

Cílem mé práce je zhodnotit dané překlady evropských legislativních textů a také najít nevhodnější překlady termínů a jejich kolokací.

Jak jsem již zmínila výše, z celkového počtu třiceti dvou klíčových slov z vybrané zkoumané oblasti „Vybrané zemědělské výrobky – tabák, len, konopí, chmel“ jsem vybrala 125 kolokací, u kterých budu následně zkoumat a analyzovat správnost českých překladů.

Pro zhodnocení správnosti překladů budu využívat zejména elektronické databáze evropských legislativních textů a to databázi ISAP<sup>15</sup> a databázi IATE<sup>16</sup>. V případě, že se mi nepodaří nalézt překlad v těchto dvou databázích, budu používat Francouzsko-český slovník, případně Česko-francouzský technický slovník uvedený v bibliografii.

Klíčová slova a jejich kolokace rozdělím do dvou hlavních skupin podle jejich překladů:

- adekvátní překlady (přesné, správné překlady) daného lexika
- neadekvátní překlady (nepřesné, chybné překlady) daného lexika

### 5.1. Adekvátní překlady

Za adekvátní, neboli přesné překlady budu považovat překlady, které jsem našla alespoň v jednom ze zdrojů (ISAP, IATE, slovníky uvedené v bibliografii, případně internet). Dále pak budu za správné (adekvátní) překlady považovat takové překlady, které mají více správných možností překladů. Vždy uvedu tabulku pro dané klíčové slovo a jeho kolokace udávající počet výskytů a překlady. Pro větší přehlednost je každá tabulka uvozena daných klíčovým slovem.

#### 5.1.1. Kolokace s jedním možným překladem

---

<sup>15</sup> ISAP = Informační systém pro aproximaci práva; dostupné na WWW: <http://isap.vlada.cz/> [cit. 1.5.10]

<sup>16</sup> IATE = Inter Active Terminology for Europe; dostupné na WWW: <http://iate.europa.eu/> [cit.1.5.10]

V této části se bude jednat o překlady bezchybné, které tudíž ve většině případů nepotřebují žádný komentář. Komentář uvedu jen tehdy, pokud bude v tabulkách uvedeno i klíčové slovo, případně když jsem musela kolokaci ověřit z jiných zdrojů.

- **accise**

*Tabulka č. 5.1.1.-1*

Kolokace	Výskytů	česky-1
accise	5	spotřební daň
accise minimale	2	minimální spotřební daň

**certification**

*Tabulka č. 5.1.1.-2*

Kolokace	Výskytů	česky-1
centres de certification	2	střediska pro ověřování
certification	77	ověřování
l'organisme de certification	2	subjekt vydávající ověřovací listiny původu
règles générales de la certification	2	obecná pravidla pro vydávání ověřovacích listin původu

- **comité**

*Tabulka č. 5.1.1.-3*

Kolokace	Výskytů	česky-1
comité	45	výbor
comité de gestion du houblon	8	Řídící výbor pro chmel
comité de gestion du lin et du chanvre	2	Řídící výbor pro len a konopí
comité de gestion du tabac	5	Řídící výbor pro tabák
Comité économique et social	13	Hospodářský a sociální výbor

Jediná nejednotnost u těchto kolokací souvisí se psaním malých a velkých písmen u francouzských názvů výborů.

- **commerce**

*Tabulka č. 5.1.1.-4*

Kolokace	Výskytů	česky-1
commerce	22	obchod



Kolokace	Výskytů	česky-1
commerce des produits	5	obchod s produkty
commerce des produits agricoles	1	obchod se zemědělskými produkty
commerce intérieur	1	vnitřní obchod
Organisation Mondiale du Commerce	1	Světová obchodní organizace

- **commercialisation**

Tabulka č. 5.1.1.-5

Kolokace	Výskytů	česky-1
campagne de commercialisation	62	hospodářský rok
commercialisation	135	uvedení na trh
commercialisation du houblon	4	uvádění chmele na trh
exigences minimales de commercialisation	11	minimální požadavky pro uvádění na trh
limites minimales de commercialisation	4	minimální požadavky pro uvedení na trh
règles de commercialisation arrêtées pour le houblon	2	opatření týkající se uvedení chmele na trh
règles relatives à la commercialisation	1	pravidla uvedení na trh
stade de commercialisation	1	fáze uvádění na trh

Klíčové slovo „*commercialisation*“ je substantivum ženského rodu, objevující se v textech celkem 135krát.

Kolokace „*règles relatives à la commercialisation*“ byla přeložena pomocí překladatelského způsobu *étouffement*. K tomuto překladatelskému postupu dochází nejčastěji při překladu do francouzského jazyka, kdy je francouzská předložka *à* doplněna plnovýznamovým slovem. V tomto případě doplněna slovem *relatif*.

Text 31971R1696

„Les *règles relatives à la commercialisation* et aux échanges avec les pays tiers s'appliquent, , en outre, aux produits suivants: ...“

„*Pravidla uvedení na trh* a pravidla obchodu se třetími zeměmi se také uplatňují na tyto výrobky: ...“

U kolokace „*limites minimales de commercialisation*“ není termínu „limite“ použito nejvhodněji, protože překlad tohoto termínu je dle VFČS „hranice“, „mez“, „rozhraní“, „pomezí“. (Neumann, J.; 1974:30)

Nicméně z kontextu je patrné, že překladatel použil termín „limite“, aby se vyhnul opětovnému použití slova „exigence“.

#### Text 31971R1696

„...présentant des caractéristiques qualitatives répondant aux *limites minimales de commercialisation* valables à un stade déterminé de la commercialisation.“

„...takových znaků jakosti, které splňují *minimální požadavky pro uvádění na trh* platné pro určenou fázi uvádění na trh.“

Po zhodnocení jsou brány tyto kolokace jako adekvátní.

- **cônes**

#### *Tabulka č. 5.1.1.-6*

Kolokace	Výskytů	Slovní druh	česky-1
houblon en cônes	26	s.m.	chmelové hlávky
cônes	36	s.m.pl.	hlávky

Anglický výraz pro chmelové hlávky je dle IATE „*hop cones*“. Toto heslo bylo v IATE hodnoceno jako spolehlivé.

Podle francouzsko-české části Francouzsko-českého a česko-francouzského slovníku je výraz „*cône du houblon*“ přeložen jako „chmelová šišťice“.

Dle skript ČZU „*chmelová hlávka (šišťice) je tedy plodenství samičí chmelové rostliny*“.<sup>17</sup>

- **contrôle**

#### *Tabulka č. 5.1.1.-7*

Kolokace	Výskytů	česky-1
contrôle	192	kontrola

<sup>17</sup>Pulkrábek J., Capouchová I. a kol.; *Speciální fyto technika*. Praha ČZU, 2003. Dostupné na WWW: [http://www3.czu.cz/php/skripta/kapitola.php?titul\\_key=4&idkapitola=181](http://www3.czu.cz/php/skripta/kapitola.php?titul_key=4&idkapitola=181) [cit. 6/5/10]

Kolokace	Výskytů	česky-1
l'organisme de contrôle	15	kontrolní subjekt

- **emballage**

*Tabulka č. 5.1.1.-8*

Kolokace	Výskytů	česky-1
emballage	7	balení
emballage final	2	konečné balení

ISAP uvádí jako překlad „*balení do dalšího obalu (tzn. vložení polotovaru v prvním obalu do dalšího obalu, např. kartonu)*“. VFČS nabízí tyto překlady: „balení“ (činnost), „obal“. Česko-francouzský technický slovník nabízí pro termín „*emballage*“ jediný správný překlad a to „balení“<sup>18</sup>.

- **feuilles**

*Tabulka č. 5.1.1.-9*

Kolokace	Výskytů	česky-1
feuilles	51	listy
tabac en feuilles	20	tabákové listy

Klíčové slovo „*feuilles*“ je podstatné jméno mužského rodu v množném čísle (sg. *feuille*), které se v textech vyskytuje celkem 51. V textech bylo toto slovo jednoznačně překládáno jako listy.

ISAP po vyhledání slova „*feuilles*“ nabízí pro oblast železa a oceli jako závazný překlad tohoto termínu slovo „*tabule*“. Ovšem dále ISAP uvádí pro další závazný termín „*feuillages, feuilles, rameaux et autres parties de plantes*“ překlad „listoví, listy, větve a ostatní části rostlin“. V tomto případě je zřejmé, že pro mou oblast výzkumu je tento překlad adekvátní.

- **fibres**

*Tabulka č. 5.1.1.-10*

<sup>18</sup> VOMÁČKOVÁ, Libuše, et al. *Česko-francouzský technický slovník*. Praha: SNTL, 1978:27

Kolokace	Výskytů	česky-1
fibres de chanvre	49	konopná vlákna
fibres	326	vlákna
du lin et du chanvre destinés à la production de fibres	19	len a konopí pěstovaný na vlákno
quantités de fibres	18	množství vláken

- **graines**

Tabulka č. 5.1.1.-11

Kolokace	Výskytů	česky-1
graines	33	semena
importations de graines de chanvre	2	dovoz semena konopí
graines de chanvre	10	semena konopí
houblon avec graines	1	chmel s peckami

Dle VFČS je slovo „graine“ překládáno jako „semeno, „zrno“. (Neumann, J.; 1974:713)

Kolokace „*houblon avec graines*“, která byla v textu přeložena jako „*chmel s peckami*“ je brána za správnou, protože Výkladový zemědělsko-potravinářský terminologický slovník<sup>19</sup> definuje termín „pecka“ jako „*vnitřní část peckovice, která obsahuje semeno (tzv. jádro)*“.

- **groupement**

Tabulka č. 5.1.1.-12

Kolokace	Výskytů	česky-1
groupement	202	seskupení
un groupement de producteurs de houblon	1	seskupení producentů chmele
groupement de producteurs	82	seskupení producentů

- **hausse**

Tabulka č. 5.1.1.-13

Kolokace	Výskytů	česky-1
----------	---------	---------

<sup>19</sup> ČAZV (Česká akademie zemědělských věd), Výkladový zemědělsko-potravinářský terminologický slovník. Dostupné na WWW: <http://www.agronavigator.cz/ts/> [cit. 4.5.10]

Kolokace	Výskytů	česky-1
hausse	18	zvýšení
hausse du prix	2	zvýšení ceny

- **houblon**

*Tabulka č. 5.1.1.-14*

Kolokace	Výskytů	česky-1
extraits végétaux de houblon	6	chmelové extrakty
houblon	378	chmel
livraison de houblon	2	dodávka chmele
produit obtenu par mouture du houblon	5	výrobek získaný mletím chmele
plante du houblon grim pant	2	rostliny chmele otáčivého

- **humidité**

*Tabulka č. 5.1.1.-15*

Kolokace	Výskytů	česky-1
taux de humidité	21	obsah vlhkosti
humidité	4	vlhkost

- **chanvre**

*Tabulka č. 5.1.1.-16*

Kolokace	Výskytů	česky-1
paille de chanvre	4	konopný stonek
chanvre	186	konopí

- **industrie**

*Tabulka č. 5.1.1.-17*

Kolokace	Výskytů	česky-1
industrie	39	průmysl
industrie du papier	4	papírenský průmysl
industrie textile	5	textilní průmysl
industrie du tabac	1	tabákový průmysl
industrie des fibres	2	odvětví vláken

- **lin**

Tabulka č. 5.1.1.-18

Kolokace	Výskytů	česky-1
secteur du lin	23	odvětví lnu
fibres courtes de lin	75	krátká lněná vlákna
superficie de lin	2	pěstební plocha lnu
pailles de lin	17	lněné stonky
lin	229	len

- **mouture**

Tabulka č. 5.1.1.-19

Kolokace	Výskytů	česky-1
mouture	5	mletí
mouture de houblon	5	mletí chmele

- **politique**

Tabulka č. 5.1.1.-20

Kolokace	Výskytů	česky-1
politique	44	politika
politique de sécurité	1	bezpečnostní politika
politique budgétaire	4	daňová politika
politique agricole commune	22	společná zemědělská politika
politique de qualité	4	politika jakosti
politique sociale	1	sociální politika
politique fiscale	2	daňová politika
politique sanitaire	2	zdravotní politika

Pro kolokaci „*politique budgétaire*“ až z kontextu vyplývá, že se jedná o daňovou politiku.

Text 52009AE0635

„1.11 Le couplage entre *politique budgétaire* et politique de santé est largement conditionné par l'établissement d'un lien entre les recettes des accises sur le tabac et les activités destinées à supprimer les conséquences négatives de la consommation de tabac. Toutefois, compte tenu du montant total des ressources financières consacrées, à l'heure actuelle, à ces activités, il est bien clair que la majeure partie des recettes des

accises sur le tabac sert à financer des activités et politiques sans aucun lien avec ces objectifs de santé publique. Il apparaît donc clairement que les objectifs budgétaires restent la priorité en ce qui concerne les droits d'accises sur les produits du tabac.“

„1.11 Vazba mezi *daňovou* a zdravotní *politikou* je do značné míry podmíněna propojením výnosů ze spotřební daně z tabáku a činností, jejichž cílem je odstranit škodlivé důsledky spotřeby tabáku. Vzhledem k celkovému financování takových činností v dnešní době je však celkem jasné, že většina výnosů ze spotřební daně z tabáku je vynakládána na činnosti a politiky, které s takovými cíli zdravotní politiky nijak nesouvisejí. Je proto zcela zřejmé, že fiskální cíle jsou stále prvořadými cíli spotřebních daní z tabákových výrobků.“

- **poudre**

*Tabulka č. 5.1.1.-21*

Kolokace	Výskytů	česky-1
poudre	28	prášek
poudre de houblon	15	chmelový prášek

- **prime**

*Tabulka č. 5.1.1.-22*

Kolokace	Výskytů	česky-1
droit à la prime	1	nárok na prémii
régime de prime	7	režim premií
partie variable de la prime	17	proměnlivá část premie
prime	129	prémie
prime au producteur	3	prémie producentovi

- **prix**

*Tabulka č. 5.1.1.-23*

Kolokace	Výskytů	česky-1
prix de vente	11	prodejní cena
prix des produits	2	cena výrobků
prix de marché	2	tržní ceny

Kolokace	Výskytů	česky-1
fixation des prix	1	stanovení cen
prix d'objectif	1	cílové ceny
prix plus élevé	4	vyšší cena
prix des cigarettes	12	cena cigaret
prix raisonnables	2	přiměřené ceny
prix moyen pondéré	4	průměrné vážená cena
évolution des prix	1	vývoj cen
fluctuations des prix	1	cenové výkyvy
prix	158	cena

Kolokace „*fluctuations des prix*“ je v textech přeložena jako „cenové výkyvy“. Ve VFČS je „*fluctuation des prix*“ přeložena jako „kolísání cen“. (Neumann, J.; 1974:656) Synonyma slova „výkyv“ jsou „odchylka“, „fluktuace“, „kolísání“. Z tohoto důvodu je „*fluctuations des prix*“ hodnocena jako adekvátní kolokace.

- **producteur**

Tabulka č. 5.1.1.-24

Kolokace	Výskytů	česky-1
producteur	131	producent
État membre producteur	4	produčenský členský stát
producteur individuel	24	samostatný producent
producteur de houblon	1	producent chmele

- **production**

Tabulka č. 5.1.1.-25

Kolokace	Výskytů	česky-1
production de fibres de chanvre	2	výroba konopných vláken
production de paille	2	produkce stonků

- **quotas**

Tabulka č. 5.1.1.-26

Kolokace	Výskytů	česky-1
réserve nationale de quotas	3	národní rezerva kvót
quotas de transformation	4	zpracovatelské kvóty
distribution de quotas	3	rozdělování kvót
répartition de quotas	1	rozdělování kvót
rachat de quotas	9	zpětný odkup kvót



Kolokace	Výskytů	česky-1
attestations de quotas de production	4	osvědčení o produkčních kvótách
quotas	112	kvóty
quotas de production	33	produkční kvóty

- **récolte**

*Tabulka č. 5.1.1.-27*

Kolokace	Výskytů	česky-1
récolte des pailles	3	sklizeň stonků
l'écoulement de la récolte	1	odbyt sklizně
l'année de récolte	11	sklizňový rok
récolte	167	sklizeň

- **superficie**

*Tabulka č. 5.1.1.-28*

Kolokace	Výskytů	česky-1
l'enregistrement des superficies	1	registrace ploch

- **tabac**

*Tabulka č. 5.1.1.-29*

Kolokace	Výskytů	česky-1
poids du tabac	2	hmotnost tabáku
Fonds communautaire du tabac	6	Tabákový fond Společenství
écoulement du tabac	2	odbyt tabáku
conservation du tabac	1	skladování tabáku
rouleaux de tabac	1	tabákové svitky
variétés de tabac	10	odrůdy tabáku
tabac	331	tabák
tabac naturel	2	surový tabák
déchets de tabac	1	tabákový odpad
tabac brut	61	surový tabák
transformation du tabac	3	zpracování tabáku
la lutte contre le tabagisme	1	boj proti závislosti na nikotinu
situation concurrentielle sur le marché du tabac	1	hospodářská soutěž na trhu s tabákem

ISAP pro „korektní hospodářskou soutěž“ nabízí závazný termín „*concurrence loyale*“ avšak po zhodnocení je kolokace brána jako adekvátní.

- **transformation**

Tabulka č. 5.1.1.-30

Kolokace	Výskytů	česky-1
contrat de transformation	10	smlouva o zpracování
transformation du houblon	6	zpracování chmele
transformation	299	zpracování

- **variétés**

Tabulka č. 5.1.1.-31

Kolokace	Výskytů	česky-1
variétés	122	odrůdy
groupe de variétés	53	skupina odrůd

### 5.1.2. Kolokace s více synonymními překlady

V následujících tabulkách uvedu překlady adekvátní, které mají více synonymních překladů. V kolonce „Výskytů“ uvedu celkový počet nalezených kolokací a v závorce pak uvedu vždy počet pro jednotlivé překlady (česky-1 + česky-2,...). Celkově do této skupiny spadá 19 kolokací a 3 klíčová slova.

- **certification**

Tabulka č. 5.1.2.-1

Kolokace	Výskytů	česky-1	česky-2	česky-3
régime de certification	5 (4+1)	režim vydávání ověřovacích listin původu	postup pro vydávání ověřovacích listin původu	
numéro de référence de certification	4 (1+3)	referenční číslo ověřovací listiny	referenční číslo ověřovací listiny původu	
procédure de certification	16 (3+5+8)	proces vydávání ověřovacích listin původu	postup pro vydávání ověřovacích listin původu	ověřovací řízení

Kolokace byly shledány jako adekvátní, protože v textech se dále skutečně mluví o listinách. Následují 2 ukázky:

Text 50005PC0386

„1. Les produits visés à l'article 1er, récoltés dans la Communauté ou élaborés à partir du houblon récolté dans la Communauté, sont soumis à une *procédure de certification*.

2. Le certificat ne peut être délivré que pour les produits présentant des caractéristiques qualitatives minimales valables à un stade déterminé de la commercialisation.“

„1. Produkty uvedené v článku 1 sklizené ve Společenství nebo získané z chmele sklizeného ve Společenství podléhají *ověřovacímu řízení* k označení původu.

2. Ověřovací listina původu může být vydána pouze pro produkty, které prokáží minimální úroveň znaků jakosti platnou pro příslušnou fázi uvádění na trh.“

#### Text 31991R2265

„6. Les représentants des organismes officiels chargés de contrôler le respect du *régime de certification* procèdent, périodiquement et sans préavis, à des contrôles d'échantillons sur les lieux-mêmes de la transformation. Ces contrôles portent sur le travail des représentants visés, l'authenticité des produits certifiés, les certificats d'accompagnement du houblon...“

„6. Zástupci úředních subjektů, kteří jsou povinni kontrolovat *režim vydávání ověřovacích listin původu*, provádějí v prostorách zpracovatelského zařízení pravidelné a neohlášené kontroly výběrových vzorků. Tyto kontroly jsou zaměřeny na práci výše uvedených zástupců, řádnost ověřených produktů, ověřovací listiny původu doprovázející chmel...“

- **commerce**

*Tabulka č. 5.1.2.-2*

Kolokace	Výskytů	česky-1	česky-2
commerce illicite	7 (4+3)	nezákonný obchod	nelegální trh

Překlad „nezákonný obchod“ se v textech vyskytuje čtyřikrát a překlad „nelegální trh“ se v textech vyskytuje třikrát.

ISAP uvádí překlad „nedovolený obchod“.

VFČS nabízí pro adjektivum „*illicite*“ tyto překlady „nedovolený“, „neoprávněný“, „zakázaný“. Dále VFČS nabízí pro kolokaci „*acte illicite*“ překlad „nedovolený čin“, „nedovolené jednání“. (Neumann, J.; 1974:763)

U překladu č. 2 není slova *trh* použito nejvhodněji, avšak z kontextu vyplývá, že překladatel chápe *trh* jako místo, kde se obchoduje, neboť „*trh*“ je prostor, kde dochází ke směně statků a peněz, čili k obchodu.

#### Text 52009AE0635

„Du point de vue d'un économiste, la hausse des droits d'accises renforce les incitations à la contrebande de cigarettes et au *commerce illicite*. La contrebande est une activité économique à laquelle s'appliquent les lois de l'offre et de la demande.“

„Z ekonomického hlediska zvýšení spotřební daně zvyšuje pobídky pro pašování cigaret a *nelegální trh*. Pašování je hospodářská činnost, pro kterou platí zákony nabídky a poptávky.“

- *côntrole*

*Tabulka č. 5.1.2.-3*

Kolokace	Výskytů	česky-1	česky-2
'attestation de contrôle	17 (16+1)	osvědčení o kontrole	kontrolní osvědčení

Překlad „osvědčení o kontrole“ se v textech vyskytl celkem 16krát, zatímco překlad „kontrolní osvědčení“ se v textech vyskytl pouze jednou.

IATE nabízí synonymní kolokaci „*certificat de contrôle*“ a k ní anglický překlad „*inspection certificate*“, pro který ISAP uvádí závazný překlad „Osvědčení o prohlídce (tj. o kontrole plavidla)“.

Ve vyhlášce č. 140/1980 Sb. federálního ministerstva zahraničního obchodu z 6. října 1980 o preventivní kontrole vyváženého a dováženého zboží a o průkazu o použití a využití dovezeného zboží nalezneme v §2 článku 5<sup>20</sup> následující: „*Kontrolní osvědčení vydá kontrolní organizace po nestranné kontrole na základě skutečně zjištěné jakosti, popřípadě zjištěného množství zboží.*“

<sup>20</sup> Vyhláška již byla zrušena, ale pro vybrané texty ji lze použít, protože se týká textů vydaných před rokem 1980.

- **emballage**

Tabulka č. 5.1.2.-4

Kolokace	Výskytů	česky-1	česky-2
unité d'emballage	3 (2+1)	balení	jednotlivé balení

Z kontextu vyplývá, že se vždy jedná o jedno balení, tím pádem se nemění smysl.

Text 31978R3076

„Chaque *unité d'emballage* faisant l'objet d'une attestation d'équivalence doit comporter les indications suivantes dans une des langues de la Communauté:...”“

„Na každém *balení*, pro které bylo vystaveno osvědčením o rovnocennosti, jsou v jednom z jazyků Společenství uvedeny tyto údaje:...”“

- **feuilles**

Tabulka č. 5.1.2.-5

Kolokace	Výskytů	česky-1	česky-2
feuilles de cône	2 (1+1)	listy hlávky	listeny hlávky

Podle encyklopedie Wikipedia<sup>21</sup> je „listen“ definován takto: „*Listen je metamorfovaný (přeměněný) list, z jehož úžlabí vyrůstají květy nebo větve květenství. Často se významně liší od podoby, kterou mají běžné fotosyntetizující listy daného rostlinného druhu.*“

- **hausse**

Tabulka č. 5.1.2.-6

Kolokace	Výskytů	česky-1	česky-2	česky-3
hausse des droits d'accises	3 (1+1+1)	zvýšení spotřební daně	nárůst spotřební daně	vyšší spotřební daně

<sup>21</sup> Dostupné na WWW: <http://cs.wikipedia.org/wiki/Listen> [cit. 12/5/10]

U překladu č. 3 bylo použito překladatelského postupu slovnědruhové transpozice, kdy francouzské substantivum bylo při překladu do češtiny nahrazeno přídavným jménem.

- ***houblon***

*Tabulka č. 5.1.2.-7*

Kolokace	Výskytů	česky-1	česky-2
superficies plantées en houblon	2 (1+1)	plochy osázené chmelem	pěstební plochy chmele
secteur du houblon	28 (3+25)	odvětví chmele	chmel

Všechny kolokace jsou synonymně překládány. Vždy je z kontextu patrné, že to stále vystihuje smysl překládaného termínu.

Text 31977R1170

„Est interdite pour la période allant du 1er juillet 1977 jusqu'au 31 décembre 1979 toute extension des *superficies plantées en houblon*.“

„V období od 1. července 1977 do 31. prosince 1979 je zakázáno jakékoli rozšiřování *pěstebních ploch chmele*.“

Text 31992R3124

„...depuis l'établissement de l'organisation commune des marchés dans le *secteur du houblon*, les échanges de produits dérivés du houblon ...“

„...že od vytvoření společné organizace trhu s chmelem se obchod s chmelovými výrobky...“

- ***champ***

*Tabulka č. 5.1.2.-8*

Kolokace	Výskytů	česky-1	česky-2	česky-3
champ	13 (8+5)	oblast	rozsah	
champ d'activité	5 (3+1+1)	oblast působnosti	rozsah činností	oblast činnosti

Ve VFČS je „*champ d'activité*“ překládáno jako „pole (oblast) působnosti“ nebo „pole (oblast) činnosti“. (Neumann, J.; 1974:292)

Text 32008R0709

„...leur *champ d'activité* et sa conformité à l'article 3,..."

„...*rozsahu* svých *činností* a souladu těchto činností s čl. 3..."

- *prime*

Tabulka č. 5.1.2.-9

Kolokace	Výskytů	česky-1	česky-2
montant de la prime	14 (2+12)	částka prémie	prémie

Všechny překlady jsou adekvátní. Vždy je z kontextu zřejmé, že se jedná o určitou částku prémie.

Text 31998R2848

„Les nouveaux prix à l'exclusion du *montant de la prime*, des services éventuels et des taxes..."

„Nové ceny bez *prémie*, poplatků za služby a daní..."

- *prix*

Tabulka č. 5.1.2.-10

Kolokace	Výskytů	česky-1	česky-2
prix d'achat	23 (21+2)	kupní cena	nákupní cena
augmentation des prix	3 (1+3)	zvýšení cen	vyšší ceny

Pro kolokaci „*prix d'achat*“ je nejpřesnější překlad č. 2, což uvádí i ISAP.

Česko-francouzský technický slovník uvádí pro tuto kolokaci obě možnosti překladu (nákupní i kupní cena). (Vomáčková, L.; 1978:52)

U kolokace „*augmentation des prix*“ není překlad č. 2 nejvhodnější, ale z kontextů vyplývá, že se jedná o zvýšení cen.

Text 52009AE0635

„Toutefois, le taux de taxation ad valorem risque aussi de décourager les producteurs d'améliorer la qualité des produits, si cela doit avoir pour conséquence une *augmentation des prix*, et donc également une augmentation de la charge fiscale.“

„Valorická sazba daně však může rovněž odrazovat výrobce od zlepšování jakosti výrobků, pokud to znamená *vyšší ceny* a tudíž i vyšší platby daně.“

- **production**

Tabulka č. 5.1.2.-11

Kolokace	Výskytů	česky-1	česky-2
régions de production	6 (2+4)	pěstitelské oblasti	produkční oblasti
production de fibres	50 (33+17)	pěstované na vlákno	výroba vláken
production	399 (358+41)	produkce	výroba

Pro kolokaci „*régions de production*“ ISAP nabízí závazný termín „produkční region“. Ale oba dva české překlady, jsou adekvátní, neboť jsou synonymní.

Pro kolokaci „*production de fibres*“ je vhodnější překlad č. 2, ale i překlad č. 1 je adekvátní, neboť z kontextu lze pochopit, že se jedná o výrobu vláken.

Text 32005R0873

„...le rapprochement des informations relatives aux parcelles agricoles cultivées en lin ou en chanvre destinées à la *production de fibres*...“

„...by měly být informace o zemědělských pozemcích, na nichž se produkuje len nebo konopí *pěstované na vlákno*,...“

- **superficie**

Tabulka č. 5.1.2.-12

Kolokace	Výskytů	česky-1	česky-2	česky-3
superficie cultivée	3 (1+1+1)	pěstební plocha	pěstitelská plocha	plocha půdy pro pěstování
superficie	40 (39+1)	plocha	velikost	



Klíčové slovo „*superficie*“ je substantivum v ženském rodě. V textech se tento termín objevoval celkem 40krát. Oba překlady jsou synonymní.

Nejpřesnějším překladem kolokace „*superficie cultivée*“ je „pěstební plocha“, což potvrzuje i ISAP. Nicméně i zbylé dva překlady jsou adekvátní. Vždy je z kontextu zřejmé, že se jedná o plochu, kde se něco pěstuje.

#### Text 52006DC0125

„... du fait que presque tous les volumes produits sont des coproduits du teillage des fibres longues et que la *superficie cultivée* spécialement pour la production de fibres courtes reste marginale.“

„... protože téměř všechny vyrobené objemy představují druhotné produkty potěrání dlouhých vláken a *plocha půdy pro pěstování* lnu určená konkrétně pro výrobu krátkých vláken zůstává marginální.“

U termínu „*superficie*“ se překlad č. 2 „velikost“ vyskytl pouze jedenkrát a to v následujícím případě:

#### Text 31998R2848

„...la *superficie* de la parcelle en cause, à l'exclusion des chemins de service et des enclos...“

„...*velikost* příslušného pozemku bez hospodářských komunikací a bez oplocení...“

- ***tabac***

*Tabulka č. 5.1.2.-13*

Kolokace	Výskytů	česky-1	česky-2
livraison du tabac	12 (5+7)	dodání tabáku	dodávka tabáku
tabac emballé	6 (4+2)	zabalený tabák	balený tabák

Jedná se o kolokace synonymní.

Pro kolokaci „*tabac emballé*“ je podle ISAPu závazný překlad „balený surový tabák“, přičemž jak jsem již zmínila výše (tabulka č. 5.1.1.-29), „surový tabák“ je překládán jako „*tabac brut*“.

## 5.2. Neadekvátní překlady

Do této skupiny jsou zařazeny kolokace, které byly alespoň jednou chybně přeloženy.

Z celkového počtu 157mi slovy a kolokacemi je tato skupina tvořena pěti kolokacemi nesprávně přeloženými.

- *comité*

Tabulka č. 5.1.3.-1

Kolokace	Výskytů	česky-1	česky-2
comité de gestion des fibres naturelles	2 (1+1)	Řídící výbor pro len a konopí	Řídící výbor pro přírodní vlákna

U této kolokace je špatně překlad č. 1 „Řídící výbor pro len a konopí“.

### Text 32002R0052

„Les mesures prévues au présent règlement sont conformes à l'avis du *comité de gestion des fibres naturelles*,...“

„Vzhledem k tomu, že opatření tohoto nařízení jsou v souladu se stanoviskem *Řídícího výboru pro len a konopí*...“

- *lin*

Tabulka č. 5.1.3.-2

Kolokace	Výskytů	česky-1	česky-2
lin textile	3 (2+1)	přadný len	lněná vlákna

Nejpřesnější překlad pro tuto kolokaci je překlad č. 1, což uvádí i ISAP.

Překlad č. 2 je špatný, neboť „vlákna“ jsou překládána jako „fibres“.

- *politique*

Tabulka č. 5.1.3.-3

Kolokace	Výskytů	česky-1	česky-2
politique de santé	6 (5+1)	zdravotní politika	oblast zdraví

#### Text 52009AE0635

„Elle laisse par ailleurs aux États membres davantage de latitude quant à l'application de droits spécifiques et au prélèvement d'accises minimales sur les cigarettes à des fins de *politique de santé* publique.“

„Navíc tento návrh umožňuje členským státům větší flexibilitu při používání zvláštních sazeb a ukládání minimálních spotřebních daní u cigaret v zájmu dosažení cílů v *oblasti zdraví*.“

- *prime*

Kolokace	Výskytů	česky-1
chasse à la prime	3	subvence

Podle encyklopedie Wikipedia<sup>22</sup> znamená „subvence“ systematickou hmotnou podporu nebo výpomoc z veřejných prostředků.

Dle VFČS „*chasse*“ znamená „lov“, „pronásledování“, „stíhání“ apod. V tomto případě by překlad byl „hon (lov) na prémie“, což neodpovídá překladu. (Neumann, J.; 1974:298)

Pro český překlad slova „subvence“ existuje francouzský termín „*subvention*“, který znamená právě „subvenci“ nebo (peněžitou) „podporu“. (Neumann, J.; 1974:570)

#### Text 52006DC0125

<sup>22</sup> Dostupné na WWW: <http://cs.wikipedia.org/wiki/Subvence> [cit. 7/5/10]

„Cette aide est octroyée aux transformateurs, les producteurs de pailles de lin et de chanvre bénéficiant, quant à eux, du régime de soutien général aux cultures arables. L'objectif recherché était d'éviter toute culture spéculative – en expansion sous le régime précédent – et de promouvoir la production de fibres longues de lin grâce au relèvement du montant de l'aide qui lui est alloué. Les résultats se sont révélés très satisfaisants: La production liée à la *chasse à la prime* a disparu, tandis que la production de lin à fibres longues a augmenté notablement.“

„Tato podpora je poskytována zpracovatelům, zatímco producenti lněných a konopných stonků využívají podporu z obecného režimu podpor pro plodiny na orné půdě. Záměrem bylo zabránit spekulativnímu pěstování, které se v rámci předcházejícího režimu rozšířilo, a podpořit výrobu dlouhých lněných vláken poskytnutím vyšší částky podpory. Výsledky jsou velmi pozitivní: vymizela výroba založená na *subvencích* a významně vzrostla výroba dlouhých lněných vláken.“

- *quotas*

*Tabulka č. 5.1.3.-4*

Kolokace	Výskytů	česky-1	česky-2
régime de quotas	8 (1+7)	režim produkčních kvót	režim kvót

Nejpřesnější překlad, pro danou kolokaci je překlad č. 2, čili „režim kvót“, což uvádí i ISAP.

V textu se objevuje i samotná kolokace „*quota de production*“, která je překlad právě pro „produkční kvótu“.

Text 31998R2848

„Les États membres constituent une base de données informatisée dans laquelle sont enregistrés, pour chacune des entreprises de transformation, pour chacun des producteurs et pour les groupements de producteurs, les indications permettant l'identification de leurs établissements ou exploitations, les quotas ou les quantités figurant sur les attestations de quota de production qui leur sont attribuées, ainsi que toute autre indication utile en vue du contrôle du *régime de quotas*.“

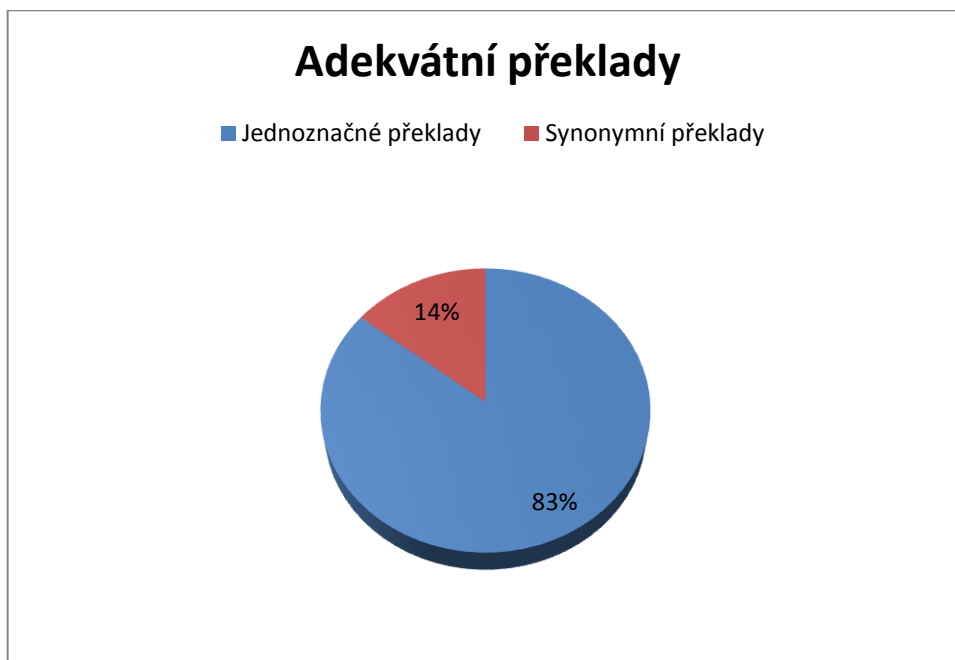
„ Členské státy vytvoří počítačovou databázi, která obsahuje o každém zpracovateli, každém samostatném producentovi a o každém seskupení producentů údaje umožňující identifikaci jejich podniků, údaje o kvótách a množstvích uvedených v osvědčeních o produkčních kvótách, jakož i všechny ostatní údaje potřebné ke kontrole *režimu produkčních kvót*.“

### 5.3. Závěr výzkumu

Během analýzy lexika nalezeného v evropských legislativních textech zaměřených na oblast: „Vybrané zemědělské produkty – tabák, len, konopí, chmel“ jsem pracovala s celkem 157mi klíčovými slovy a kolokacemi. Z tohoto celkového počtu bylo celkem 5 kolokací přeložených neadekvátně, což představuje pouze 3% z celkového počtu mého zkoumaného lexika (viz. *Graf č. 5: Adekvátnost překladů*).

Zbýlých 152 termínů a kolokací je přeloženo adekvátně, což představuje 97% z celkového zkoumaného lexika (viz *Graf č. 4: Adekvátní překlady*). V některých případech se jednalo o překlady adekvátní s jedním jednoznačným překladem – 130 klíčových slov a kolokací, která představují 83% z celkového počtu zkoumaného lexika. Druhou skupinu adekvátních překladů tvoří 22 klíčových slov a kolokací, které byly přeloženy více synonymními překlady, což čítá 14% z celkového počtu klíčových slov a kolokací.

*Graf č. 4: Adekvátní překlady*



Graf č. 5



## 6. Závěr

Výzkum této práce vychází z vytvoření paralelního korpusu. Tento paralelní korpus je sestaven ze 30-ti evropských legislativních textů ve francouzském a českém jazyce z období let 1971 – 2009. Česká verze obsahuje 70 473 slov a francouzská verze 90 394 slov. Všechny texty jsou zaměřeny na oblast „Vybrané zemědělské výrobky – tabák, len, konopí chmel.“ Na základě tohoto korpusu jsem sestavila databázi čítající 32 klíčových slov a 125 kolokací.

Tato práce je rozdělena do dvou hlavních částí. V první, teoretické části jsem se podrobněji zabývala výběrem textů a popisem jeho úprav. Dále jsem popsala práci s mým korpusem, který obsahuje 30 textů, a tyto texty dávají dohromady 274,4 stran v českém jazyce. Z tohoto základu jsem si následně vybrala 157 klíčových slov a jejich kolokací, které byly předmětem zkoumání a jsou z hlediska vztahu k dané oblasti zajímavé.

Hlavním cílem této práce bylo vyhodnocení adekvátnosti překladů daných klíčových slov a kolokací. Co se týče kvality vyhodnocení překladu klíčových slov a kolokací, tak celkem 152 klíčových slov a jejich kolokací bylo hodnoceno jako správně přeložené, což je 97%. Zbýlých 5 kolokací, tedy 3% bylo vyhodnoceno jako nesprávně přeložené. Jedná se o následující kolokace: „*régime de quotas*“, „*chasse à la prime*“, „*politique de santé*“, „*comité de gestion des fibres naturelles*“, „*lin textile*“.

Druhým cílem této bakalářské práce bylo vytvoření tematického francouzsko-českého slovníčku, který obsahuje vybraná klíčová slova a jejich spojení k dané oblasti „Vybrané zemědělské výrobky – tabák, len, konopí chmel.“



## 7. Résumé

Ce mémoire de licence est basé sur le travail de recherche d'un corpus parallèle. Ce corpus parallèle est composé de 30 textes législatifs européens en langue française et tchèque dans les années 1971 - 2009. La version tchèque contient 70.473 mots et la version française - 90.394 mots. Tous les textes sont orientés sur le domaine «Produits agricoles sélectionnés - tabac, lin, chanvre, houblon.»

La première partie de ce travail est focalisée sur la sélection et présentation de 30 textes sélectionnés en français et en tchèque. Les textes choisis comptent 316,3 pages dans la langue française et en tchèque 274,4 pages. La deuxième partie décrit comment modifier le texte et travailler avec le programme ParaConc que nous avons créé un corpus parallèle sur la base des textes législatifs choisis. Egalement dans cette section j'ai décrit la sélection de mots-clés et les collocations. J'ai choisi 32 mots-clés (termes) et puis j'ai trouvé leurs 125 collocations. Ensuite, dans cette section consacrée toujours à expliquer le concept de collocation et les types de collocations.

Le troisième chapitre est consacré à des explications de concepts linguistiques de corpus, qui est un des concepts les plus récents et dont le développement a contribué au développement des ordinateurs et des technologies de l'informatique. J'ai aussi décrit quels types de corpus existent.

Enfin, la partie théorique est consacrée à la classification des structures morphologiques des collocations. J'ai classé des collocations en 8 catégories selon la classification créée par Mathieu-Colas. La classe de collocations qui était la plus représentée dans ma recherche a été la classe numéro 12. La structure N+de+X compte 70 collocations qui représente 56% du nombre total des collocations.

Dans la partie pratique, j'ai enfin analysé l'adéquation des traductions de mots-clés et de leurs collocations ce qui était l'objectif principal de ce mémoire de licence. Quant à l'évaluation de la qualité des traductions de mots-clés et collocations je l'ai divisée en 2 catégories principales - les traductions adéquates et des traductions inadéquates. Les traductions adéquates sont en plus divisées en deux parties selon le nombre d'équivalents des traductions tchèques pour une collocation – en traductions traduites d'une manière univoque et en celles qui sont traduites par plusieurs processus synonymes.

De nombre total 157 mots-clés et les collocations, j'ai trouvé que 152 traductions ont été évaluées comme correctement traduites donc adéquates, ce qui représente 97% du nombre total. Pour des collocations inadéquates ce n'est que 3%, c'est-à-dire seulement 5 collocations examinées mal traduites.

Le deuxième objectif de ce mémoire de licence était la création d'un dictionnaire français-tchèque thématiques. Ce dictionnaire comprend les mots-clés sélectionnés et leurs collocations spécialisés dans la domaine «Produits agricoles sélectionnés - tabac, lin, chanvre, houblon.» Ce dictionnaire fait partie de l'annexe.

## 8. Seznam použité literatury

### Seznam odborné literatury:

Čermák F.; *Jazyk a jazykověda*, Nakladatelství Karolinum, Praha, 2004

Blatná R., Čermák F.; *Korpusová lingvistika: Stav a modelové přístupy*, Nakladatelství Lidové noviny, Praha, 2006, ISBN 80-7106-865-9

Čermák F., Šulc M.; *Studie z korpusové lingvistiky – Kolokace*, Nakladatelství Lidové noviny, Praha, 2006

Čermák F., Klímová J., Petkevič V.; *Studie z korpusové lingvistiky*, Nakladatelství Karolinum, Praha, 2000

Čermák F., Schmiedtová V.; *Český národní korpus: základní charakteristika a širší souvislosti*, In: Národní knihovna: knihovnická revue č. 3., 2004  
Dostupné na: <http://knihovna.nkp.cz/pdf/0403/0403152> [cit. 27.4.10]

Blatná R.; *Korpus – uskutečněný sen*, In: Tvar, 1997, č. 9

Čermák F., *Syntagmatika slovníku: typy lexikálních kombinací*, In: Čeština - univerzália a specifika 3, eds. Z. Hladká, P. Karlík, Brno, 2001  
Dostupné na: <https://mathesius.ff.cuni.cz/ucnk/doc/synkol.rtf> [cit. 15.4.10]

Mathieu-Colas, M., *Essai de typologie des noms composés français*. In: Cahiers de Lexicologie, 69, Paris: Didier Erudition, 1996  
Dostupné na: [http://halshs.archives-ouvertes.fr/halshs-00433841\\_v1/](http://halshs.archives-ouvertes.fr/halshs-00433841_v1/) [staženo 20.4.10]

Blatná R.; *Jazyková databanka neboli korpus*, In: Vesmír, 1997, č. 12 str. 670-671  
Dostupné na: [www.vesmir.cz/files/file/fid/199/aid/3118](http://www.vesmir.cz/files/file/fid/199/aid/3118) [cit. 27.4.10]

Pulkrábek J.; Capouchová I. a kol; *Speciální fyto technika*, ČZU, Praha, 2003  
Dostupné na: [http://www3.czu.cz/php/skripta/kapitola.php?titul\\_key=4&idkapitola=181](http://www3.czu.cz/php/skripta/kapitola.php?titul_key=4&idkapitola=181) [cit. 6.5.10]

### Slovníky:

Neumann J., Hořejší V. a kol. *Velký francouzsko-český slovník I. (A – K)*, 1.vyd., Academia, Praha, 1974

Neumann J., Hořejší V. a kol. *Velký francouzsko-český slovník II. (L – Z)*, 1.vyd., Academia, Praha, 1974

Vomáčková L. a kol.; *Francouzsko-česky technický slovník*, SNTL – Nakladatelství technické literatury, Praha, 1978

Petráčková V., Kraus J. a kol.; *Akademický slovník cizích slov I. (A – K)*, 1. vyd., Academia, Praha, 1995

Petráčková V., Kraus J. a kol.; *Akademický slovník cizích slov II. (L – Ž)*, 1. vyd., Academia, Praha, 1995

Vlasák V.; *Francouzsko-český a česko-francouzský slovník, část francouzsko-česká*, 1. vyd., SPN, Praha, 1997

Kolektiv autorů; *Francouzsko-český česko-francouzský studijní slovník*, 2. vyd., FIN PUBLISHING, Praha, 2005

Elektronické zdroje:

*Interactive Terminology for Europe* dostupné na <http://iate.europa.eu>

*Informační systém pro aproximaci práva* dostupné na <http://isap.vlada.cz>

*Vyhláška o preventivní kontrole vyvážení a dovážení zboží a o průkazu o použití a využití dovezeného zboží* dostupné na [http://www.lexdata.cz/lexdata/sb\\_free.nsf/c12571cc00341df1000000000000000/c12571cc00341df1c12566d40072cfdc?OpenDocument](http://www.lexdata.cz/lexdata/sb_free.nsf/c12571cc00341df1000000000000000/c12571cc00341df1c12566d40072cfdc?OpenDocument)

*Paraconc* dostupné na [www.mt-archive.info/LREC-2002-Barlow.pdf](http://www.mt-archive.info/LREC-2002-Barlow.pdf)

*EUR-Lex Přístup k právu Evropské unie* dostupné na <http://eur-lex.europa.eu/cs/index.htm>

## 9. Přílohy

### Seznam použitých zkratk

s.m. – substantivum maskulinum

s.f. – substantivum femininum

pl. – plurál

sg. – singulár

N – Nom (podstatné jméno)

Adj. – Adjective (přídavné jméno)

SN - syntagme nominal (jmenné syntagma)

Det. – Determinant (člen)

X – nahrazuje jiný znak

ISAP – Informační systém pro aproximaci práva

IATE – Inter-Active Terminology for Europe

VFČS – Velký francouzsko-český slovník

ČZU – Česká zemědělská univerzita

## Francouzsko-český slovníček termínů a kolokací

Slovníček je složen ze 156 klíčových slov a kolokací, které se v této práci vyskytují. Jsou uvedeny všechny adekvátní překlady a u čtyř nesprávně přeložených kolokací je uveden správný překlad. Pouze není uvedena kolokace „*chasse à la prime*“, která je nesprávně přeložena (správný termín by byl „*subvention*“, který by tak odpovídal českému překladu „subvence“).

**ACCISE (s.f.)** *spotřební daň*

**accise minimale** *minimální spotřební daň*

**CERTIFICATION (s.f.)** *ověřování*

**centres de certification** *střediska pro ověřování*

**numéro de référence de certification** *referenční číslo ověřovací listiny; referenční číslo ověřovací listiny původu*

**organisme de certification** *subjekt vydávající ověřovací listiny původu*

**procédure de certification** *ověřovací řízení k označení původu; proces ověřovacích listin původu; ověřovací řízení*

**régime de certification** *režim vydávání ověřovacích listin původu; postup pro vydávání ověřovacích listin původu*

**règles générales de la certification** *obecná pravidla pro vydávání ověřovacích listin původu*

**COMITÉ (s.m.)** *výbor*

**comité de gestion des fibres naturelles** *Řídícího výboru pro přírodní vlákna*

**comité de gestion du houblon** *Řídící výbor pro chmel*

**comité de gestion du lin et du chanvre** *Řídící výbor pro len a konopí*

**comité de gestion du tabac** *Řídící výbor pro tabák*

**Comité économique et social** *Hospodářský a sociální výbor*

**COMMERCE (s.m.)** *obchod*

**commerce des produits agricoles** *obchod se zemědělskými produkty*

**commerce des produits** *obchod s produkty*

**commerce illicite** *nezákonný obchod; nelegální trh*

**commerce intérieur** *vnitřní obchod*

**Organisation Mondiale du Commerce** *Světová obchodní organizace*

**COMMERCIALISATION (s.f.)** *vedení na trh*

**campagne de commercialisation** *hospodářský rok*

**commercialisation du houblon** *uvádění chmele na trh*

**exigences minimales de commercialisation** *minimální požadavky pro uvádění na trh*

**limites minimales de commercialisation** *minimální požadavky pro uvedení na trh*

**règles de commercialisation arrêtées pour le houblon** *opatření týkající se uvedení chmele na trh*

**règles relatives à la commercialisation** *pravidla uvedení na trh*

**stade de commercialisation** *fáze uvádění na trh*

**CÔNES (s.m.pl.)** *hlávky*

**houblon en cônes** *chmelové hlávky*

**CONTRÔLE (s.m.)** *kontrola*

**attestation de contrôle** *osvědčení o kontrole; kontrolní osvědčení*

**organisme de contrôle** *kontrolní subjekt*

**EMBALLAGE (s.m.)** *balení*

**emballage final** *konečné balení*

**unité d'emballage** *balení; jednotlivé balení*

**FEUILLES (s.m.pl.)** *listy*

**feuilles de cône** *listy hlávky; listeny hlávky*

**tabac en feuilles** *tabákové listy*

**FIBRES (s.f.pl.)** *vlákna*

**du lin et du chanvre destinés à la production de fibres** *len a konopí pěstovaný na vlákno*

**fibres de chanvre** *konopná vlákna*

**quantités de fibres** *množství vláken*

**GRAINES (s.f.pl.)** *semena*

**graines de chanvre** *semena konopí*

**houblon avec graines** *chmel s peckami*

**importations de graines de chanvre** *dovozy semen konopí*

**GROUPEMENT (s.m.)** *seskupení*

**groupement de producteurs de houblon** *seskupení producentů chmele*

**groupement de producteurs** *seskupení producentů*

**HAUSSE (s.f.)** *zvýšení*

**hausse des droits d'accises** *zvýšení spotřební daně; nárůst spotřební daně; vyšší spotřební daně*

**hausse du prix** *zvýšení ceny*

**HOUBLON (s.m.)** *chmel*

**extraits végétaux de houblon** *chmelové extrakty*

**livraison de houblon** *dodávka chmele*

**plante du houblon grim pant** *rostliny chmele otáčivého*

**produit obtenu par mouture du houblon** výrobek získaný mletím chmele  
**secteur du houblon** odvětví chmele; chmel  
**superficies plantées en houblon** plochy osázené chmelem; pěstební plochy chmele

**HUMIDITÉ (s.f.)** vlhkost  
**taux de humidité** obsah vlhkosti

**CHAMP (s.m.)** oblast; rozsah  
**champ d'activité** rozsah činnosti; oblast působnosti; oblast činnosti

**CHANVRE (s.m.)** konopí  
**paille de chanvre** konopný stonek

**INDUSTRIE (s.m.)** průmysl  
**industrie des fibres** odvětví vláken  
**industrie du papier** papírenský průmysl  
**industrie du tabac** tabákový průmysl  
**industrie textile** textilní průmysl

**LIN (s.m.)** len  
**fibres courtes de lin** krátká lněná vlákna  
**lin textile** přadný len  
**pailles de lin** lněné stonky  
**secteur du lin** odvětví lnu  
**superficie de lin** pěstební plocha lnu

**MOUTURE (s.f.)** mletí  
**mouture de houblon** mletí chmele

**POLITIQUE (s.f.)** politika  
**politique agricole commune** společná zemědělská politika  
**politique budgétaire** daňová politika  
**politique de qualité** politika jakosti  
**politique de santé** zdravotní politika  
**politique de sécurité** bezpečnostní politika  
**politique fiscale** daňová politika  
**politique sanitace** zdravotní politika  
**politique sociale** sociální politika

**POUDRE (s.m.)** prášek  
**poudre de houblon** chmelový prášek

**PRIME (s.f.)** prémie



**droit à la prime** *nárok na prémii*  
**montant de la prime** *částka prémie; prémie*  
**partie variable de la prime** *proměnlivá část prémie*  
**prime au producteur** *prémie producentovi*  
**régime de prime** *režim premií*

**PRIX (s.m.)** *cena*  
**augmentation des prix** *zvýšení cen; vyšší ceny*  
**évolution des prix** *vývoj cen*  
**fixation des prix** *stanovení cen*  
**fluctuations des prix** *cenové výkyvy*  
**prix d'achat** *kupní cena; nákupní cena*  
**prix de marché** *tržní ceny*  
**prix de vente** *prodejní cena*  
**prix des cigarettes** *cena cigaret*  
**prix des produits** *cena výrobků*  
**prix d'objectif** *cílové ceny*  
**prix moyen pondéré** *průměrné vážená cena*  
**prix plus élevé** *vyšší cena*  
**prix raisonnables** *přiměřené ceny*

**PRODUCTEUR (s.m.)** *producent*  
**État membre producteur** *producentův členský stát*  
**producteur de houblon** *producent chmele*  
**producteur individuel** *samostatný producent*

**PRODUCTION (s.f.)** *produkce; výroba*  
**production de fibres de chanvre** *výroba konopných vláken*  
**production de fibres** *pěstované na vlákno; výroba vláken*  
**production de paille** *produkce stonků*  
**régions de production** *pěstitelské oblasti; produkční oblasti*

**QUOTAS (s.m.pl.)** *kvóty*  
**attestations de quotas de production** *osvědčení o produkčních kvótách*  
**distribution de quotas** *rozdělování kvót*  
**quotas de production** *produkční kvóty*  
**quotas de transformation** *zpracovatelské kvóty*  
**rachat de quotas** *zpětný odkup kvót*  
**régime de quotas** *režim kvót*  
**répartition de quotas** *rozdělování kvót*  
**réserve nationale de quotas** *národní rezerva kvót*

**RÉCOLTE (s.f.)** *sklizeň*  
**année de récolte** *sklizňový rok*

**écoulement de la récolte** *odbyt sklizně*  
**récolte des pailles** *sklizeň stonků*

**SUPERFICIE (s.f.)** *plocha; velikost*  
**enregistrement des superficies** *registrace ploch*  
**superficie cultivée** *pěstební plocha; pěstitelská plocha; plocha půdy*

**TABAC (s.m.)** *tabák*  
**conservation du tabac** *skladování tabáku*  
**déchets de tabac** *tabákový odpad*  
**écoulement du tabac** *odbyt tabáku*  
**Fonds communautaire du tabac** *Tabákový fond Společenství*  
**livraison du tabac** *dodání tabáku; dodávka tabáku*  
**lutte contre le tragisme** *boj proti závislosti na nikotinu*  
**poids du tabac** *hmotnost tabáku*  
**rouleaux de tabac** *tabákové svitky*  
**situation concurrentielle sur le marché du tabac** *hospodářská soutěž na trhu s tabákem*  
**tabac brut** *surový tabák*  
**tabac emballé** *balený tabák; zabalený tabák*  
**tabac naturel** *surový tabák*  
**transformation du tabac** *zpracování tabáku*  
**variétés de tabac** *odrůdy tabáku*

**TRANSFORMATION (s.f.)** *zpracování*  
**contrat de transformation** *smlouva o zpracování*  
**transformation du houblon** *zpracování chmele*

**VARIÉTÉS (s.f.pl.)** *odrůdy*  
**groupe de variétés** *skupina odrůd*