

Ciclo de Conferências UNILEGIS – 2009
Brasília, 29-30 junho

Terminologia, Lingüística de *Corpus*
e Organização da Informação Legislativa e Jurídica

Anna Maria Becker Maciel

Grupo TERMISUL

PPG Letras, Instituto de Letras

UFRGS



Qual é o ponto de encontro da Terminologia, da Lingüística de *Corpus* e da Organização da Informação Legislativa e Jurídica?

1 Terminologia

1.1 Terminologias e Terminologia

1.2 Panorama histórico

1.2.1 Século XX: Wüster e a TGT

1.2.2 Século XXI: novas propostas

1.3 Linguagem jurídica

1.3.1 Terminologia Jurídica

1.3.2 Termo jurídico

2 Lingüística de *Corpus*

2.1 Bases teórico- metodológicas

2.2 Instrumental

2.2.1 *Corpus*

2.2.2 Ferramentas básicas

2.3 Exemplos de exploração de *corpus*

3 Terminologia, Lingüística de *Corpus* e a Organização da Informação Legislativa e Jurídica

TERMINOLOGIA

- terminologia

Conjunto de palavras temática e pragmaticamente marcadas utilizadas na comunicação da informação de uma área especializada.

Ex: terminologia da nanotecnologia
terminologia das mudanças climáticas
terminologia do artesanato de lã de ovinos
terminologia da física quântica
terminologia do cajú
terminologia da moda, do carnaval, do futebol

TERMINOLOGIA

- Terminologia

Estudo sistemático da **denominação dos conceitos** utilizados na **comunicação do conhecimento** em **áreas especializadas do saber e da experiência humana** quando consideradas no seu funcionamento social.

- Objeto: **o termo**
- Objetivo: **a comunicação e a fixação do conhecimento especializado**

Panorama histórico

- **primórdios:** 1º registro **2340 a.C** : sumérios, acadianos
- **séc. XVII:** Furetière **DICTIONAIRE UNIVERSEL** Contenant généralement tous les mots François tant vieux que modernes, & les termes de toutes les Sciences & des Arts (1690)
- **séc. XVIII**
Lineu: nomenclatura binomial *Systema naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis* (1735)

Diderot et D'Alembert *Encyclopédie ou dictionnaire raisonné des sciences, des arts et des métiers* (1750):

Lavoisier, Bertholet, Fourcroy nomenclatura química em francês (1788)
- **séc. XIX:** Sociedade Técnica Russa

Panorama histórico (continuação)

- 1930 – 1970: **Origem e estruturação**
 - Lotte: Escola Soviética de Terminologia
 - Wüster: Escola de Viena (TGT)
- 1970 – 1990: **Expansão**
 - Europa, Canadá, América Latina
- 1990 – 2000: **Novas propostas**
 - Socioterminologia
 - Terminologia Textual
 - Teoria Comunicativa da Terminologia
 - Teoria Cognitiva da Terminologia
- 2000 – 2009: **Consolidação**

Wüster (1898-1974) e a Teoria Geral da Terminologia (TGT)

- *Internationale Sprachnormung in der Technik, besonders in der Elektrotechnik.* (1931) A normalização Internacional na Técnica, especialmente na Eletrotécnica
- *The machine tool: An interlingual dictionary of basic concepts.* (1968). A máquina ferramenta: um dicionário interlíngüe de conceitos básicos
- *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie.* (1979) Introdução à teoria geral da Terminologia e lexicografia terminológica

Die Allgemeine Terminologielehre — ein Grenzgebiet zwischen Sprachwissenschaft, Logik, Ontologie, Informatik und den Sachwissenschaften

(WÜSTER, 1974, p.62)

A **Ontologia** e a **Lógica** constroem a perspectiva em que a **Terminologia** estuda e organiza a denominação dos conceitos das ciências que estudam os objetos - Física, Eletrotécnica, a Medicina, a Sociologia e o Direito - (WÜSTER, 1974, p.103).

A **Lingüística** fornece o instrumental que torna possível a visualização do termo como um signo no sistema da língua.

Enquanto a **Informática** e a **Documentação** se encarregam do armazenamento, tratamento e recuperação da informação

Pressupostos da TGT

- O **conceito** é um construto mental, elaborado a partir da síntese das características de fenômenos do mundo real ou imaginário empiricamente apreendidos, devidamente despojados de marcas individualizantes.
- O **conceito** é identificado por um símbolo, o signo lingüístico, posicionado, de acordo com os traços temáticos que carrega, na estrutura de conhecimento da área temática, a árvore de domínio.
- Na árvore de domínio, o **termo** é unívoco, conciso, livre de conotação e ambigüidades e monorreferencial.

Terminologia e o cenário socio-cultural das duas últimas décadas

- Multiplicidade e diversidade de contextos comunicativos: línguas, culturas, políticas, situações econômicas e culturais.
- Crescimento acelerado da ciência e da técnica.
- Globalização do conhecimento.
- Disponibilidade de recursos informatizados.
- Avanços na ciência e na tecnologia da informação.
- Avanços dos estudos lingüísticos.

Estudos lingüísticos

Principais avanços que influenciaram as mudanças no paradigma clássico da Terminologia

- Revalorização do **componente léxico**
- Teorias de **base cognitiva**
- Lingüística **sistêmico-funcional**
- Lingüística **textual**
- Análise do **discurso**
- Lingüística de ***corpus***

Novas teorias propostas

- Socioterminologia (GAUDIN, 1993)
- Teoria Comunicativa da Terminologia (CABRÉ, 1999)
- Terminologia Textual (SLODZIAN, 2000).
- Teoria Sociocognitiva da Terminologia (TEMMERMAN, 2000)

Perspectiva comum das novas propostas

- Valorização do contexto cognitivo e sociolingüístico do termo
- Valorização do texto como o evento comunicativo real
- Valorização do usuário das terminologias

Pressupostos básicos das novas propostas

- A **conceitualização** não pode ser separada da atividade da fala, tampouco pode ser uma atividade isolada e individual, posto que é fruto de uma construção que se realiza no uso efetivo da língua em um âmbito especializado.
- O **termo** é uma unidade léxica de valor pragmático ativado no processo de comunicação pela sua função de vetor de informação de uma área temática.
- O **termo** é uma unidade léxica da língua comum cuja especificidade é ativada pela conjugação do seu caráter de representante de um conceito temático e suas funções pragmáticas assumidas na comunicação.

Terminologia do século XXI

Quais são os rumos da terminologia do futuro:

- Um campo interdisciplinar para o qual convergem três teorias principais: Teoria do Conhecimento, Teoria da Comunicação, Teoria da Linguagem.
- Adequação e harmonização da metodologia descritiva e dos objetivos normalizadores, de acordo com as necessidades do usuário visado e os propósitos da aplicação terminológica a ser elaborada.

Linguagem Jurídica

- Não há uma língua própria do Direito, há uma utilização da **língua vernácula** com especificidades do universo jurídico.
- O Direito fala por muitas bocas, sobre muitos temas, em muitos lugares, abrangendo três grandes dimensões principais — doutrina, legislação e jurisprudência — que imprimem na **linguagem jurídica** características próprias de acordo com propósitos e contextualizações peculiares.

termo

- “[...] **palavra** ativada singularmente por suas condições pragmáticas de adequação a um tipo de comunicação” (Cabré, 1999, p.123).
- “[...] as **unidades lexicais** só se tornam **termos** quando são definidas e utilizadas nos textos de especialidade” (Kokourek, 1991, p.105).
- **Unidade léxica** de caráter **denominativo e conceitual** que denomina um núcleo na estrutura conceitual de uma área temática.

Terminologia jurídica

Tudo o que o Direito julga necessário para que a ordem social seja garantida, torna-se jurídico no momento em que é criada uma norma legal para circunscrever seu conceito e regulamentar sua existência.

“A incidência da regra jurídica é que torna jurídicos os bens da vida” (PONTES de MIRANDA, 1970).

A palavra *casa* passa a ser uma unidade terminológica jurídica, quando contemplada na perspectiva do Direito Constitucional e do Direito Penal

- a Constituição assegura ao cidadão o *direito da inviolabilidade* de sua *casa* no Art. 5º XI;
- o Art. 150 § 4º do Código Penal define *casa*;
- *casa*, na legislação brasileira, é um termo jurídico do sistema de direitos e garantias fundamentais do cidadão.

O termo jurídico se caracteriza pelo aspecto conceitual temático unido à funcionalidade pragmática

- Ao significado usual de *pai, mãe, cônjuge, família, casamento, férias, empregado, empregador, salário, benefício*, acrescentam-se, no **universo jurídico**, traços específicos que vinculam essas entidades à esfera cognitiva e funcional da estrutura organizada de direitos, deveres e obrigações estabelecidas pela lei.

Termos jurídicos por excelência

*anticrese,
decurus,
exequatur,
exequenda,
enfiteuse,
sucumbência,
habeas data.*

Termos jurídicos de dupla pertinência

Termos jurídicos na língua comum

*usofruto,
herança,
salvo-conduto,
quorum,
tribunal, juiz,
crime,
julgamento.*

Língua comum + sentido jurídico

*despejo,
servidão,
julgado,
imposto,
sentença,
testador,
tombamento,
tempestivo,
falecer.*

Termos com o sentido comum e implicações legais

*pai, mãe, filho,
família, parente,
cônjuge,
empregado,
férias, salário,
benefício.*

Termos jurídicos na linguagem especializada e na língua comum

(categorização inspirada em CORNU, G. *Linguistique juridique*. Paris: Montchrestien, 1990. p. 61-87).

Especificidade do termo jurídico

- A **terminologia jurídica** se configura através do uso no contexto sociocultural da lei.
- As características temáticas reveladas pelo significado semântico são insuficientes para caracterizar o **termo jurídico**.
- A especificidade do **termo jurídico** é ativada pela implicação pragmática do seu referente com os outros referentes na estrutura deôntica do Direito.
- A inserção no **universo jurídico** do referente se dá através das suas implicações pragmáticas com outros referentes na estrutura deôntica da área.

Lingüística de Corpus

lingüística com *corpus* ?

Lingüística de Corpus?

Lingüística Computacional?

Corpus

No Dicionário eletrônico Houaiss da língua portuguesa 1.0

- *Corpus* substantivo masculino
- **1** coletânea ou conjunto de documentos sobre determinado tema
- Ex.: ele estuda o c. juris canonici (coletânea de direito canônico)
- **2** Derivação: por analogia. repertório ou conjunto da obra científica, técnica e/ou artística de uma pessoa ou a ela atribuída; Ex.: o c. da poética camoniana
- **3** Rubrica: fisiologia; estrutura com características ou funções especiais no corpo de um homem ou de um animal
- **4** Rubrica: **lingüística**; conjunto de enunciados numa determinada língua, ger. colhidos de atos reais da fala, que servem como material para análise lingüística
- **5** Rubrica: **lingüística, semiologia**; conjunto de enunciados (que são indefinidamente possíveis, i.é., inesgotáveis), constituído por amostras significativas da gramática de determinada língua. Ex.: o c. sintagmático da língua portuguesa
- pl.: *corpòra* ; emprt. lat. *còrpus nom. sing. de corpus,òris 'corpo'*
- emprt. lat. *còrpus nom. sing. de corpus,òris 'corpo'*

O que se entende por *corpus* na Lingüística de *Corpus*

- Uma coletânea de textos autênticos
- compilados segundo critérios pré-estabelecidos
- completos ou na forma de excertos
- de linguagem natural
- escritos ou orais
- produzidos sem a intervenção do pesquisador
- produzidos para comunicar informações e não para exemplificar o uso da língua
- legíveis pelo computador
- preparados para a pesquisa lingüística.

O que **não é um corpus** na **Linguística de Corpus**

- Um conjunto de palavras
- Um conjunto de frases
- Um arquivo de textos
- Um conjunto de arquivos de textos
- Um dicionário ou um conjunto de dicionários
- Uma base de dados
- Uma base de conhecimento
- Uma biblioteca
- *A WEB*

ainda que em formato digital

Usuários *de corpora* informatizados

- Pesquisadores cujo objeto de estudo é a língua, especialmente lexicógrafos, lexicólogos, terminólogos e terminógrafos.
- Tradutores
- Redatores Técnicos
- Professores de língua: materna, estrangeira, instrumental
- Gestores da informação
- Bibliotecários, Documentalistas, Arquivistas, Indexadores
- Informatas: Processadores da linguagem Natural (PLN); Desenvolvedores de *software* para Tradução Automática (TA); Especialistas em Inteligência Artificial (IA); Mineração da Informação; Web Semântica, Construção de Ontologias.

Os usuários de *corpora* informatizados fazem Lingüística de *Corpus* ou Lingüística Computacional?

- Não necessariamente.
- Pode-se fazer o estudo científico da língua utilizando *corpora* e não fazer nem Lingüística Computacional nem Lingüística de *Corpus*.
- O *corpus* não é o objeto de estudo de uma área lingüística, é uma ferramenta, o utensílio com que se faz o estudo. Todas as áreas que têm a língua como objeto de estudo podem usar um *corpus* informatizado como ferramenta.

Lingüística de *Corpus* ≠ Lingüística Computacional

- A Lingüística de *Corpus* enfoca a **comunicação** dos falantes humanos entre si.
- A Lingüística Computacional visa à **comunicação** do homem com a máquina.

Lingüística de *Corpus* ≠ Lingüística Computacional

Lingüística de *Corpus*:

- área das ciências humanas que estuda o uso da língua em grandes conjuntos de textos autênticos através de **evidências empíricas** extraídas por meio de programas de computador;
- dá primazia ao **desempenho** do falante em contextos reais de comunicação, prioriza a visão **sociolingüística**, a **configuração sistêmico-funcional** do significado e o caráter **probabilístico** da linguagem;
- não ignora a intuição do pesquisador;
- harmoniza pesquisa qualitativa e quantitativa.

Lingüística de *Corpus* ≠ Lingüística Computacional

Lingüística Computacional:

- área da ciência da computação com base nas ciências exatas que estuda a formalização da língua escrita e falada com vistas ao desenvolvimento de técnicas informatizadas adequadas para o seu processamento, interpretação e geração;
- vê a língua como um conjunto de regras finitas e aproxima-se das teorias lingüísticas racionalistas que dão primazia á competência do falante e concede prioridade ao processamento da língua natural (PLN).

Pressupostos básicos da Lingüística de *Corpus*

- valorização da observação empírica em grandes grandes extensões de textos
 - o uso concreto da linguagem em contextos reais e interlocutores autênticos
- concepção funcionalista da linguagem
 - o significado se configura no contexto de uso
- visão da língua como um sistema probabilístico
 - a seleção léxica e funcional da combinação de palavras forma padrões recursivos

Padrões léxicos: **perdas e danos** no Código Civil (brCCpt TERMISUL 89.201 palavras)

seu exercício sofrer limitação voluntária Art 12 Pode se exigir que cesse a ameaça ou a lesão a direito da personalidade e reclamar **perdas** e danos sem prejuízo de outras sanções previstas em lei Parágrafo único Em se tratando de morto terá legitimação para requerer a medida prevista neste

145 São os negócios jurídicos anuláveis por dolo quando este for a sua causa Art 146 O dolo accidental só obriga à satisfação das **perdas** e danos e é accidental quando a seu despeito o negócio seria realizado embora por outro modo Art 147 Nos negócios jurídicos

se a parte a quem aproveite dele tivesse ou devesse ter conhecimento em caso contrário ainda que subsista o negócio jurídico o terceiro responderá por todas as **perdas** e danos da parte a quem ludibriou Art 149 O dolo do representante legal de uma das partes só obriga o representado a responder civilmente até

representado a responder civilmente até a importância do proveito que teve se porém o dolo for do representante convencional o representado responderá solidariamente com ele por **perdas** e danos Art 150 Se ambas as partes procederem com dolo nenhuma pode alegá lo para anular o negócio ou reclamar indenização

Vicia o negócio jurídico a coação exercida por terceiro se dela tivesse ou devesse ter conhecimento a parte a que aproveite e esta responderá solidariamente com aquele por **perdas** e danos Art 155 Subsistirá o negócio jurídico se a coação decorrer de terceiro sem que a parte a que aproveite dela tivesse ou

se a coação decorrer de terceiro sem que a parte a que aproveite dela tivesse ou devesse ter conhecimento mas o autor da coação responderá por todas as **perdas** e danos que houver causado ao coacto SEÇÃO IV Do Estado de Perigo Art 156 Configura se o estado de perigo quando alguém

Padrões Léxicos: **anticrese** no Código Civil (brCCpt TERMISUL 89.201 palavras) colocados: **penhor, hipoteca, bens**

objeto atribuir direitos preferenciais mediante hipoteca penhor ou **anticrese** sua invalidade importará somente na anulação da preferência ajustada

ao credor ressaltar o penhor a hipoteca ou a **anticrese** se os bens dados em garantia pertencerem a terceiro

a hipoteca X a **anticrese** Art 1 226 Os direitos reais

TÍTULO X Do Penhor da Hipoteca e da **Anticrese** CAPÍTULO I Disposições Gerais Art 1

1 419 Nas dívidas garantidas por penhor **anticrese** ou hipoteca o bem dado em garantia fica sujeito

que pode alienar poderá empenhar hipotecar ou dar em **anticrese** só os bens que se podem alienar poderão ser

que se podem alienar poderão ser dados em penhor **anticrese** ou hipoteca § 1 o A propriedade superveniente

Art 1 424 Os contratos de penhor **anticrese** ou hipoteca declararão sob pena de não terem eficácia

arrematação ou da adjudicação CAPÍTULO IV Da **Anticrese** Art 1 506 Pode o devedor ou

imputado ao capital § 2 o Quando a **anticrese** recair sobre bem imóvel este poderá ser hipotecado pelo

assim como o imóvel hipotecado poderá ser dado em **anticrese** Art 1 507 O credor anticrético

Padrões semânticos: **cominar** prosódia negativa (LgBrMA – TERMISUL 340.743 palavras)

disciplinar a que a lei **comine** a pena de demissão ou

incide nas penas a estes **cominadas** considerados em especial os

incidirá nas penas a eles **cominadas** 6º Se o autor

Justiça cópia da decisão **cominativa** da pena aplicada no

demolição sem prejuízo da **cominação** de outras penalidades previstas em

A imposição de contrapropaganda será **cominada** quando o fornecedor incorrer na

incide nas penas a esses **cominadas** na medida de sua culpabilidade

da pena privativa da liberdade **cominada** ao crime Na individualização

às exigências legais mediante **cominações** que terá eficácia de

à metade do valor máximo **cominado** em abstrato §

Padrão sintático (Recorte das 76 ocorrências no corpus LgBrMA – TERMISUL 340.743 palavras)

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o CONGRESSO NACIONAL decreta e eu sanciono a

○ PRESIDENTE DA REPÚBLICA **faço** saber que o Congresso Nacional decreta e eu sanciono a

Metodologia interativa: homem+máquina

Máquina

- Instrumental
 - *corpus*
 - ferramentas básicas
 - listador
 - concordanciador
 - gerador de N-gramas
 - outras ferramentas
 - etiquetador
 - alinhador de textos
 - concordanciador alinhado

Usuário

- Tratamento do *corpus*
 - compilação
 - exploração
 - processamento
 - coleta de dados
 - descrição
 - análise e discussão
- interpretação
- validação dos dados
 - procedimentos estatísticos

Listador de palavras: Lista **individualmente todas as palavras** presentes no texto e/ou no *corpus* por ordem alfabética ou de frequência. (brCopt TERMISUL46.656 palavras)

abandonado	1
abastecimento	1
aberta	1
abertos	3
abertura	10
abolir	1
abono	3
abrang	2
abrangendo	2
abrangerá	2
abrangidas	3
abrangidos	3
absoluta	29

Listagem p/ordem alfabética: **Poder** ocorre 120 vezes na
(brCopt TERMISUL46.656 palavras)

pluralismo	2
pluralista	1
plurianuais	2
plurianual	15
pluripartidarismo	1
pobres	1
pobreza	2
pode	8
podem	9
podendo	26
poder	120

Listagem p/ordem de freqüência: Poder

brCopt TERMISUL46.656 palavras)

público	185
iii	185
nacional	184
união	177
pelo	173
aos	172
estado	170
sobre	167
tribunal	163
sua	151
forma	151
estados	145
presidente	142
distrito	142
república	134
será	129
iv	124
poder	120

Concordanciador

Lista as ocorrências no texto de uma determinada palavra com seu contexto imediato. (recorte de 120 ocorrências de **poder** na Constituição 88 (brCopt TERMISUL46.656 palavras))

Parágrafo único Todo o **poder** emana do povo que o exerce por meio de

em defesa de direitos ou contra ilegalidade ou abuso de **poder**

a lei não excluirá da apreciação do **Poder** Judiciário lesão ou ameaça a direito

sua liberdade de locomoção por ilegalidade ou abuso de **poder**

data quando o responsável pela ilegalidade ou abuso de **poder** for autoridade pública ou agente de pessoa jurídica no exercício

ou agente de pessoa jurídica no exercício de atribuições do **poder** público

previdência social com reajustes periódicos que lhe preservem o **poder** aquisitivo sendo vedada sua vinculação para qualquer fim

ressalvado o registro no órgão competente vedadas ao **poder** público a interferência e a intervenção na organização sindical

a normalidade e legitimidade das eleições contra a influência do **poder** econômico ou o abuso do exercício de função cargo

Recorte das 110 ocorrências de **Justiça** na Constituição 88 (brCopt TERMISUL46.656 palavras)

a igualdade e a **justiça** como valores supremos de uma

poderá ser impugnado ante a **Justiça** Eleitoral no prazo de quinze

mandato tramitará em segredo de **justiça** respondendo o autor

prestação de contas à **Justiça** Eleitoral

Prefeito perante o Tribunal de **Justiça**

o Tribunal de **Justiça** der provimento a representação para

do Superior Tribunal de **Justiça** ou do Tribunal Superior Eleitoral

dos Desembargadores do Tribunal de **Justiça** limitado a noventa inteiros

Desembargadores do respectivo Tribunal de **Justiça** limitado a noventa inteiros

membros do Conselho Nacional de **Justiça** e do Conselho Nacional do

quando o decretar a **Justiça** Eleitoral nos casos previstos

Ministros do Superior Tribunal de **Justiça** aplicando se lhes

o Ministro da **Justiça**

o Ministro da **Justiça**

Gerador de N-gramas Lista as ocorrências de grupos de palavras repetidos ao longo de um texto ou *corpus* com indicação do seu número de freqüência. **brCopt** (TERMISUL 46.656 palavras)

Total de N-Gramas: 34083

na forma da lei	79
do distrito federal e	42
o disposto no art	38
nos termos da lei	35
estados do distrito federal	32
dos estados do distrito	32
distrito federal e dos	31
pelo presidente da república	27
do supremo tribunal federal	26
a que se refere	24
de que trata o	24
e do distrito federal	23
do presidente da república	23
que se refere o	22
federal e dos municípios	20
membros do ministério público	19
distrito federal e os	19
o presidente da república	19
o presidente da república	18

Busca de termos (LgBrMA – TERMISUL 340.743 palavras)

de recursos hídricos § 2º A declaração de **reserva** de disponibilidade hídrica será transformada automaticamente pelo respectivo poder

de energia hidráulica § 3º A declaração de **reserva** de disponibilidade hídrica obedecerá ao disposto no art 13

a emissão temporária pela ANELL das declarações de **reserva** de disponibilidade hídrica de que trata o art 7º

I Estação Ecológica II **Reserva** Biológica III Parque Nacional

um mil e quinhentos hectares Art 10 A **Reserva** Biológica tem como objetivo a preservação integral da biota e

e os processos ecológicos naturais § 1º A **Reserva** Biológica é de posse e domínio público sendo que

III Floresta Nacional IV **Reserva** Extrativista V Reserva de Fauna

IV Reserva Extrativista V **Reserva** de Fauna VI Reserva de Desenvolvimento Sustentável

V Reserva de Fauna VI **Reserva** de Desenvolvimento Sustentável VII Reserva Particular

Reserva de Desenvolvimento Sustentável VII **Reserva** Particular do Patrimônio Natural Art 15 A Área

Contextos definitórios (LgBrMA – TERMISUL 340.743 palavras)

ao juro de seis por cento ao ano Parágrafo único Poderão acumular o empréstimo de que trata este artigo dois ou mais trabalhadores rurais que se **entenderem** para aquisição de propriedade de área superior à que estabelece o número 2 do artigo 4º desta Lei sob a administração comum ou em forma de cooperativa

Lei determinará as condições e exigências a serem cumpridas para licenciamento dos manipuladores de gêneros alimentícios
TÍTULO V Notificação Compulsória Art 54 Para efeito desta Lei **entende** se por notificação compulsória a comunicação à autoridade sanitária de casos confirmados ou suspeitos das doenças que por sua gravidade incidência ou possibilidade de disseminação exijam medidas

de cujo reconhecimento depende a sua regularização sem prejuízo da fiscalização específica do Imposto de Renda Artigo 4º Para os fins da presente lei **entende** se como despesas de florestamento e reflorestamento aquelas que forem aplicadas diretamente pelo contribuinte ou mediante a contratação de serviços de terceiros na elaboração do projeto técnico no preparo

inclusive a educação da comunidade objetivando capacitá-la para participação ativa na defesa do meio ambiente art 3 Para os fins previstos nesta Lei **entende** se por I meio ambiente o conjunto de condições leis influências e interações de ordem física química e biológica que permite

a problemática ambiental e social IV o incentivo à participação individual e coletiva permanente e responsável na preservação do equilíbrio do meio ambiente **entendendo** se a defesa da qualidade ambiental como um valor inseparável do exercício da cidadania V o estímulo à cooperação entre as diversas regiões do País

a problemática ambiental e social IV o incentivo à participação individual e coletiva permanente e responsável na preservação do equilíbrio do meio ambiente **entendendo** se a defesa da qualidade ambiental como um valor inseparável do exercício da cidadania V o estímulo à cooperação entre as diversas regiões do País

Lista alfabética: busca de paradigmas morfológicos (LgBrMA – TERMISUL 340.743 palavras)

cometerem	1
cometida	5
cometidas	8
cometido	4
cometidos	3
cometimento	9
cometimentos	1
cominação	2
cominações	3
cominada	3
cominadas	9
cominado	2
cominativa	2
comine	1
cominuição	1
comissão	170
comissionados	3
comissionamento	1

Webcorp output for query: cominação Domain: br
Recorte da página <http://www.webcorp.org.uk/>

- <http://www.tj.ro.gov.br/emeron/sapem/2001/setembro/2809/ARTIGOS/A08.htm>
[Plain Text](#) [Word List](#) Código Brasileiro de Trânsito). A cominação da pena pode ser especial
- http://www.mundojuridico.adv.br/sis_artigos/artigos.asp?codigo=691
[Plain Text](#) [Word List](#) defina, nem pena sem prévia cominação legal'. Da mesma forma seguiu
- <http://www.cvm.gov.br/port/descol/resp.asp?File=2006-026D04072006.htm>
[Plain Text](#) [Word List](#) poderia ter sido solicitada sob cominação de multa. Eventual investigação, caso
- http://www.tre-ro.gov.br/trenet/eleicao2006/colet_legislacao/acordaos/2002/Acor330_02-1040_02.htm
[Plain Text](#) [Word List](#) da propaganda. Irrelevância para a cominação final. Propaganda eleitoral irregular em

Etiquetador: insere no *corpus* automática ou semi-automáticamente códigos que categorizam as palavras.

- A etiquetagem pode ser:
 - a) morfológica (substantivo, adjetivo, etc.);
 - b) sintática (sujeito, predicado, etc)
 - c) semântica (ex: casa=moradia; martelo=ferramenta)
 - d)discursiva (referentes anafóricos, tópicos, etc.)

Início do preâmbulo da Constituição 88

nós [nós] **PERS** M/F 1P NOM @SUBJ> #1->82
representantes [representante] **N** M/F P @N<PRED #3->1
de [de] <sam-> **PRP** @N< #4->3
o [o] <artd> <-sam> **DET** M S @>N #5->6
povo [povo] **N** M S @P< #6->4
brasileiro [brasileiro] **ADJ** M S @N< #7->6
,reunidos [reunir] **V** PCP M P @N<PRED #9->6
em [em] **PRP** @<ADV L #10->0
Assembléia=Nacional=Constituinte
[Assembléia=Nacional=Constituinte] **PROP** M S @P< #11->10
para [para] **PRP** @ADV L> #12->82
instituir [instituir] **V** INF @ICL-P< #13->12
um [um] <arti> **DET** M S @>N #14->15
Estado=Democrático [Estado=Democrático] **PROP** M S
@<ACC

World of visl Portuguese automatic parsing:

<http://visl.sdu.dk/visl/pt/parsing/automatic/>

Concordanciador alinhado

[ENG] 1 - Man is both creature and moulder of his environment, which gives him physical sustenance and affords him the opportunity for intellectual, moral, social and spiritual growth. In the long and tortuous evolution of the human race on this planet a stage has been reached when, through the rapid acceleration of science and technology, man has acquired the power to transform his **environment** in countless ways and on an unprecedented scale. Both aspects of man's environment, the natural and the man-made, are essential to his well-being and to the enjoyment of basic human rights the right to life itself.

[PORT] 1 - O homem é ao mesmo tempo criatura e criador do meio ambiente, que lhe dá sustento físico e lhe oferece a oportunidade de desenvolver-se intelectual, moral, social e espiritualmente. A longa e difícil evolução da raça humana no planeta levou-a a um estágio em que, com o rápido progresso da Ciência e da Tecnologia, conquistou o poder de transformar de inúmeras maneiras e em escala sem precedentes o meio ambiente. Natural ou criado pelo homem, é o meio ambiente essencial para o bem-estar e para gozo dos direitos humanos fundamentais, até mesmo o direito à própria vida.

[ENG] 2 - The protection and improvement of the human **environment** is a major issue which affects the well-being of peoples and economic development throughout the world; it is the urgent desire of the peoples of the whole world and the duty of all Governments.

[PORT] 2 - A proteção e a melhoria do meio ambiente humano constituem desejo premente dos povos do globo e dever de todos os Governos, por constituírem o aspecto mais relevante que afeta o bem-estar dos povos e o desenvolvimento do mundo inteiro.

[ENG] 5 - The natural growth of population continuously presents problems for the preservation of the environment, and adequate policies and measures should be adopted, as appropriate, to face these problems. Of all things in the world, people are the most precious. It is the people that propel social progress, create social wealth, develop science and technology and, through their hard work, continuously transform the human environment. Along with social progress and the advance of production, science and technology, the capability of man to improve the **environment** increases with each passing day.

[PORT] 5 - O crescimento natural da população suscita a toda hora problemas na preservação do meio ambiente, mas políticas e medidas adequadas podem resolver tais problemas. De tudo o que há no mundo, a associação humana é o que existe de mais preciosa. É ela que impulsiona o progresso social e cria a riqueza, desenvolve a Ciência e a Tecnologia e, através de seu trabalho árduo, continuamente transforma o meio ambiente. Com o progresso social e os avanços da produção, da Ciência e da Tecnologia, a capacidade do homem para melhorar o meio ambiente aumenta dia a dia.

Corpora em língua portuguesa e Ferramentas de livre acesso on line

- TERMISUL: <http://www6.ufrgs.br/termisul>
- TEXTQUIM: <http://www6.ufrgs.br/textquim/index.php>
- COMET: <http://www.fflch.usp.br/dlm/comet/>
- LácioWeb: <http://www.nilc.icmc.usp.br>
- Linguateca: <http://www.linguateca.pt>
- <http://www.corpusdoportugues.org/x.asp>
- CEPRIIL: lael.pucsp.br/corpora
- Unitex: www-igm.univ-mlv.fr/~unitex/
- *WEBasCorpus: <http://www.webcorp.org.uk>*
(*Não é um corpus, é uma ferramenta auxiliar*)

Conclusão

- A interface entre a Organização da Informação Legislativa e Jurídica, a Terminologia e a Lingüística de *Corpus* é a palavra que significa um conceito na linguagem jurídica.
- A **linguagem jurídica** é o ponto de encontro das três áreas, cada uma com suas especificidades, propósitos, objetos de estudo e metodologias próprias.

Referências (citadas nesta apresentação ppt)

CABRÉ, M.T. (1999) *La terminología: representación y comunicación: elementos para una teoría de base comunicativa e otros artículos*. Barcelona, IULA.

CORNU, Gérard. (1990) *Linguistique juridique*. Paris: Monchrestien, 1990

GAUDIN, François. (1993) *Pour une socioterminologie. Des problèmes sémantiques aux pratiques institutionnelles*. Rouen: Publications de l'Université de Rouen. 231p.

KOCOUREK, Rostislav (1991) *La langue française de la technique et de la science*. Wiesbaden: Oscar Brandstetter Verlag GMBH.

SLODZIAN, M. (2000). L'émergence d'une terminologie textuelle et le retour du sens. In: BÉJOINT, H. ;Thoiron, P. (orgs.). *Les senses in terminologie*. Lyon: Presses Universitaires de Lyon. p 61-85.

Referências (citadas nesta apresentação ppt) [continuação]

TEMMERMANN, R. (2000). *Towards new ways of terminology description. The sociocognitive approach*. Amsterdam/New York: John Benjamins.

WÜSTER, Eugen. (1974) Die allgemeine Terminologielehre - ein Grenzgebiet zwischen Sprachwissenschaft, Logik, Ontologie, Informatik und den Sachwissenschaften. *Linguistics*, n.119. La Hague: Mouton. p.61-106.

Contato: annamb.maciel@gmail.com

Para saber mais sobre

• pesquisa terminológica com *corpus* acesse **Educação continuada** em <http://www6.ufrgs.br/termisul/>

• **Lingüística de Corpus**: acesse [http://pt.wikipedia.org/wiki/Linguistica de Corpus](http://pt.wikipedia.org/wiki/Linguistica_de_Corpus)

Leia:

Berber Sardinha, T. (2004). *Lingüística de Corpus*. São Paulo: Manole.