

## FICÇÃO, TRADUÇÃO, TERMINOGRAFIA E LINGUÍSTICA DE *CORPUS*: CONFLUÊNCIAS

Guilherme FROMM

Universidade Federal de Uberlândia  
[guifromm@ileel.ufu.br](mailto:guifromm@ileel.ufu.br)

**Resumo:** o objetivo deste trabalho é mostrar como alunos de Iniciação Científica podem ser treinados para trabalhar na área de Terminografia Bilíngue a partir de projetos individuais. Para tanto, descrevemos os passos para a elaboração de um dicionário técnico na área de ficção através de exemplos compilados de *corpora* de legendas de séries televisivas disponíveis na Internet e sua posterior visualização, disponível no site do VoTec.

**Palavras-chave:** Ficção; Terminografia Bilíngue; Linguística de Corpus; Iniciação Científica.

### 1. Introdução

O objetivo deste trabalho é mostrar como treinar alunos de Iniciação Científica para trabalharem com a área de Terminografia Bilíngue. Normalmente, em grupos de pesquisa, os alunos de Iniciação coletam e trabalham com dados referentes aos projetos de seus orientadores. No nosso caso, embora o projeto de pesquisa original do orientador trate da compilação de *corpora* para a construção de terminologias científicas, um novo projeto foi criado especialmente para esse treinamento dos orientandos: Terminologia e Ficção. A ideia é que os alunos aprendam todo o processo de levantamento e análise de dados de forma individual, desenvolvendo um subprojeto completo dentro desse projeto maior.

O tema encadeador para que todos os alunos envolvidos nesse projeto de treinamento possam trabalhar as terminologias disponíveis na ficção são as séries de televisão que oferecem legendas em línguas inglesa e portuguesa disponíveis na Internet. Essas legendas são coletadas, processadas através da metodologia da Linguística de *Corpus* e os dados inseridos num banco de dados terminográfico, o VoTec<sup>1</sup> (Vocabulário Técnico Online; FROMM, 2007). A página de consulta desse banco está disponível na Internet (<http://www.ic.voteconline.com.br>) e o acesso é gratuito.

### 2. A escolha do tema

Ao tentar fugir do padrão de pesquisadores trabalhando em grupo através de grandes projetos em áreas diversas do conhecimento para a construção de terminologias, mono- ou bilíngues (os projetos TERMISUL, disponível em

---

<sup>1</sup> O projeto do VoTec propõe que os termos ali inseridos e todas as informações disponibilizadas na página de consulta sejam construídos, única- e exclusivamente, a partir dos dados coletados nos *corpora* de especialidade.

<http://www6.ufrgs.br/termisul/index.php> e [GETERM](http://www.geterm.ufscar.br/), disponível em <http://www.geterm.ufscar.br/>, são exemplos claros dessa tendência), a ideia de trabalhar individualmente com a ficção surgiu a partir de conversas entre o orientador e seus orientandos de Iniciação Científica, todos integrantes do Grupo de Pesquisa em Linguística de Corpus do ILEEL/UFU. A área de Terminologia, que nasceu a partir das propostas de Wüster para normalização de vocabulário de áreas técnicas (BARROS, 2006), normalmente trabalha com a padronização dos termos de determinadas áreas de especialidade. Os vocábulos ou termos (seguindo a proposta de BARBOSA, 2001), no entanto, já há algum tempo, saíram das áreas de especialidade para frequentarem o falar cotidiano de grandes parcelas da população.

Entre os programas de televisão norte-americanos, nota-se, desde a década de 1960, que séries de ficção científica avançam no uso de termos emprestados de várias áreas do conhecimento. Para este projeto, três tipos de padrões no uso de terminologia em séries televisivas foram identificados:

a. séries com terminologia totalmente ficcional: figuram neste item os programas sobre o fantástico, com personagens e características que existem apenas em mundos criados a partir da imaginação. Exemplos clássicos são as séries sobre monstros, fantasmas, espíritos, mitologias, etc. No nosso caso, a série estudada foi *Supernatural* (CARNEIRO, 2011);

b. séries que misturam ficção e ciência: campo propício para a ficção científica, nesses programas encontramos terminologia própria de áreas reais como Física, Química, Astronáutica, Astronomia, Engenharias, Biologia, etc. No entanto, como a ficção também se faz presente, termos são criados, dentro dessas áreas, para explicar fenômenos que ainda não podem ser explicados ou, ainda, fenômenos inventados. Uma das séries mais famosas nesse campo e que deve ser trabalhada dentro do projeto é *Star Trek*; pesquisamos, também, a série australiana *Farscape* (SILVA, 2011). Essas séries, assim como as do item c, contam com consultores (MARTHE, 2011), advindos dessas respectivas áreas técnicas, que fornecem termos reais (ou, pelo menos, com “aura” de reais) para melhor situarem os enredos;

c. séries que retratam o cotidiano de médicos, investigadores, cientistas forenses, etc.: nesse caso, as terminologias usadas tentam retratar, de forma fiel, àquelas usadas pelos profissionais dessas áreas nos seus afazeres cotidianos. Podemos destacar, entre outras, *House* (BANG, 2011), *CSI* (LAGO, 2011), *Law and Order* e *Grey's Anatomy*, sendo que as duas primeiras estão entre os nossos objetos de pesquisa.

Após a sugestão de trabalho com uma série por parte do orientando e o aval do orientador<sup>2</sup>, inicia-se a fase de compilação dos *corpora* em línguas inglesa e portuguesa e a posterior análise dos mesmos, via programas de Análise Lexical.

### 3. Metodologia: o levantamento de legendas

Embora haja muita controvérsia sobre a disponibilização, por parte de legendadores alternativos<sup>3</sup> (SYURI, 2011), de legendas gratuitas na Internet, não levamos essas discussões em consideração. O que nos importa é o material linguístico

<sup>2</sup> É bastante comum o orientando achar que qualquer série televisiva serviria para o projeto. Muitos propõem comédias para a análise. Cabe ao orientador, através de seu conhecimento de terminologia e da série proposta, ou através de análises prévias com ferramentas de Análise Lexical, verificar se essa série atende às especificidades do projeto.

<sup>3</sup> No caso das legendas em português, os tradutores nem sempre são profissionais da área de Tradução. São profissionais de várias áreas que fazem o trabalho por hobby, sem nenhum tipo de pagamento.

disponibilizado para análise. As legendas no original são criadas, normalmente, a partir do recurso *Closed Caption* (para deficientes auditivos). Vários sites disponibilizam (como o [www.legendas.tv](http://www.legendas.tv)), em várias línguas, as traduções dos originais. Em questão de minutos, várias temporadas das séries em análise podem ser baixadas. A vantagem, em termos de trabalho com programas de Análise Lexical, é que essas legendas já vêm no formato .txt (ou .srt, facilmente convertidas em .txt), melhor lido por esses programas. Na figura 1, verificamos um diretório com os arquivos da primeira temporada da série *House* e as legendas do primeiro episódio; na figura 2, um exemplo dos arquivos em inglês e português do primeiro episódio de *Enterprise* (uma das cinco séries de *Star Trek*).

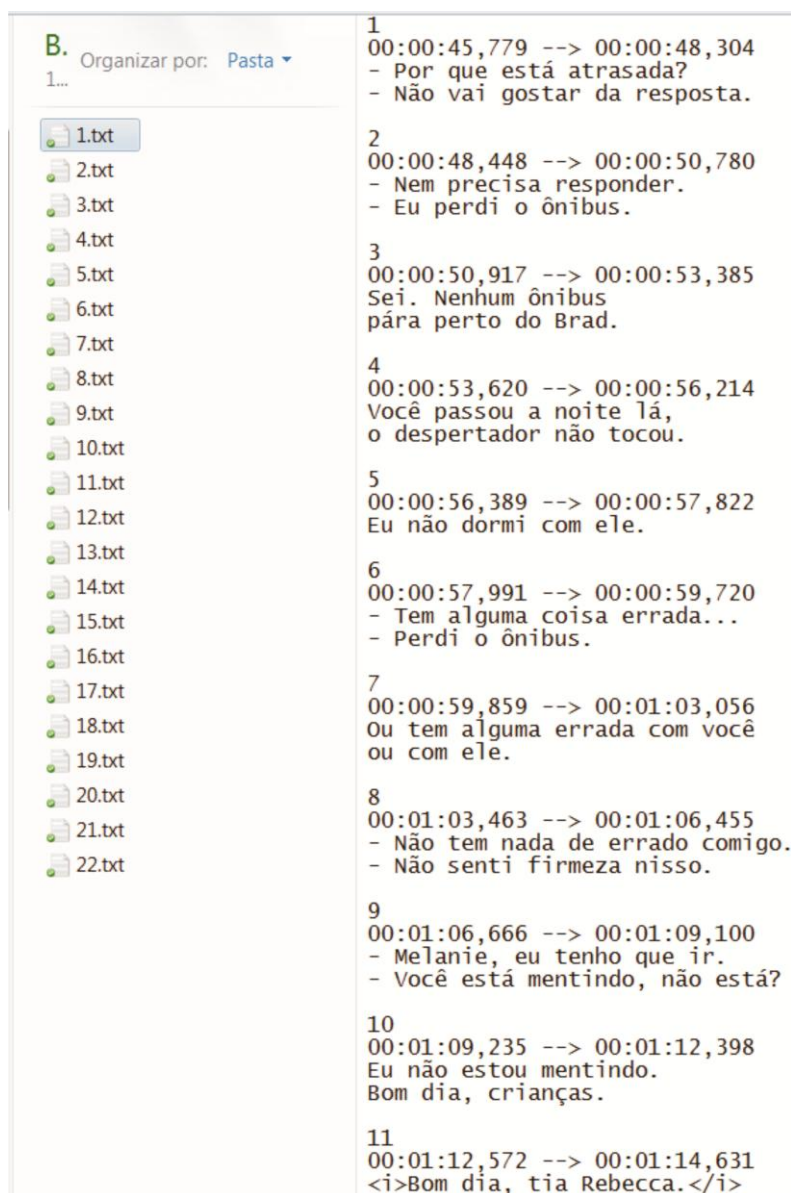


Figura 1. Diretório da primeira temporada de *House* e exemplo das legendas do primeiro episódio.

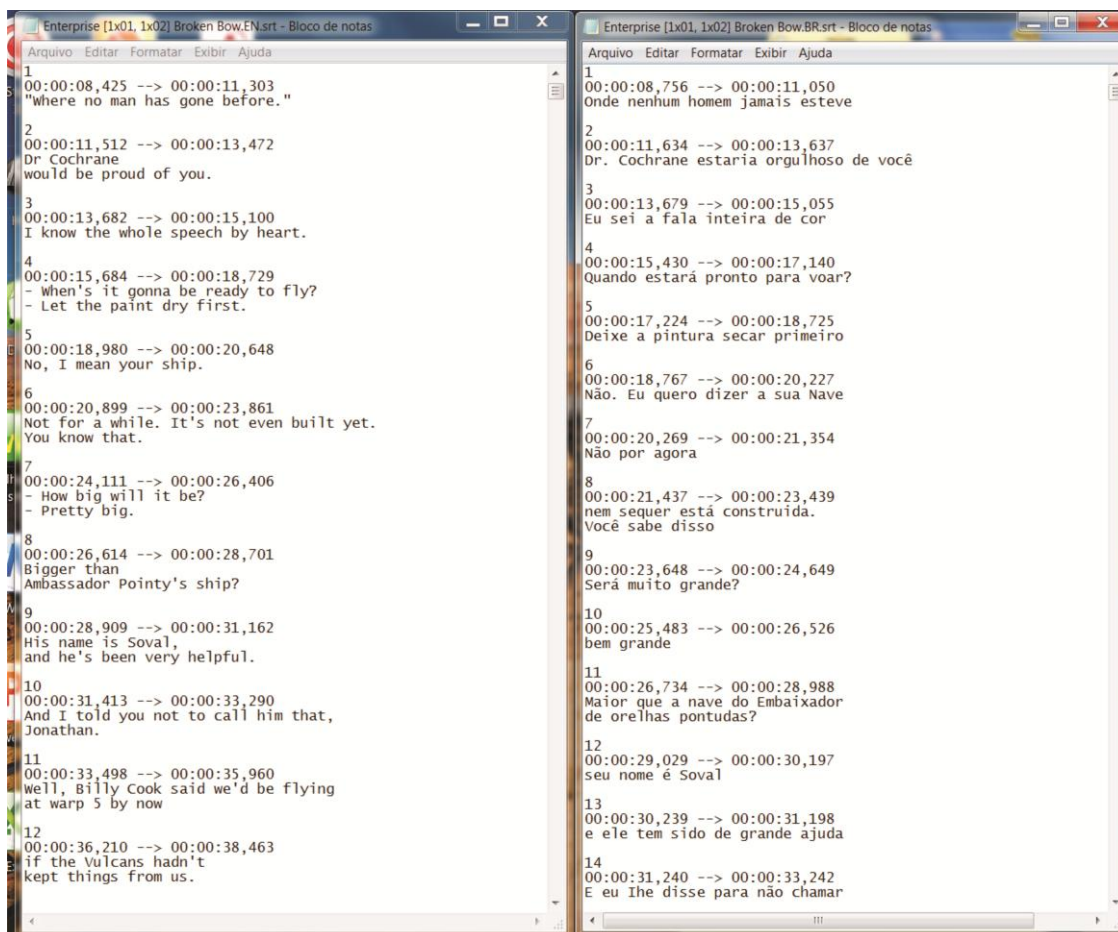


Figura 2. Exemplo das legendas, em inglês e português, do primeiro episódio de *Enterprise* (*Star Trek*).

A constituição dos *corpora* deste projeto, seguindo os parâmetros propostos por Berber Sardinha (2004), é a seguinte:

- a. Modos: falados (transcrições do seriado original para o inglês) e escritos (traduções);
- b. tempos: sincrônicos e contemporâneos;
- c. seleção: por amostragem (estático);
- d. balanceado: as legendas das séries são analisadas na íntegra, com todas as temporadas disponíveis, tanto no original quanto na tradução;
- e. conteúdos: especializados, multilíngues;
- f. autoria: de língua nativa (falantes nativos);
- g. disposições internas: paralelos (original e tradução);
- h. finalidades: de estudo (*corpus* a ser descrito).

#### 4. Metodologia: a análise dos *corpora*

A análise do material coletado é feita através do programa de Análise Lexical *WordSmith Tools* (SCOTT, 2008) e suas três ferramentas: *Wordlist*, *Keywords* e *Concordance*. Descrevemos cada uma a seguir:

- a. *Wordlist*: a ferramenta permite elaborar uma listagem com todas as palavras únicas (*types*) em relação ao total de palavras (*tokens*) do *corpus* de estudo selecionado. Na figura 3a, por exemplo, temos uma listagem que compreende as legendas de dez anos da

série *CSI* em português, já com um filtro<sup>4</sup> aplicado. Podemos notar, no lado inferior esquerdo, que o total de *types* da série é de 34.204 palavras. Na figura 3b, na parte superior direita, percebemos que o tamanho total do corpus é de 2.167.191 palavras (*tokens – running words in text*), embora só 892.346 (*tokens used for wordlist*) tenham sido usadas na análise, já que o programa ignora os horários das legendas;

Figure 3a. Wordlist da série *CSI*.

N	Word	Freq	%	Texts	%
1	E	15,257	0.70	232	100.00
2	ESTÁ	6,413	0.30	232	100.00
3	FOI	4,959	0.23	232	100.00
4	ESTAVA	3,873	0.18	232	100.00
5	TEM	3,672	0.17	231	99.57
6	ERA	2,809	0.13	231	99.57
7	QUANDO	2,407	0.11	232	100.00
8	PODE	2,277	0.11	232	100.00
9	TER	2,009	0.09	232	100.00
10	OU	2,004	0.09	231	99.57
11	SER	1,925	0.09	232	100.00
12	CARA	1,884	0.09	224	96.55
13	HÁ	1,871	0.09	229	98.71
14	VAI	1,869	0.09	232	100.00
15	SABE	1,706	0.08	232	100.00
16	TINHA	1,705	0.08	227	97.84
17	ESTOU	1,689	0.08	230	99.14
18	VOU	1,676	0.08	231	99.57
19	SÃO	1,673	0.08	232	100.00
20	SEI	1,630	0.08	231	99.57
21	CASA	1,619	0.07	220	94.83
22	DISSE	1,589	0.07	230	99.14
23	SANGUE	1,545	0.07	216	93.10
24	ACHO	1,530	0.07	229	98.71
25	CARRO	1,500	0.07	203	87.50
26	LÁ	1,499	0.07	231	99.57
27	TENHO	1,488	0.07	231	99.57
28	PARECE	1,485	0.07	230	99.14
29	COISA	1,472	0.07	228	98.28
30	FAZER	1,390	0.06	231	99.57
31	QUER	1,344	0.06	228	98.28
32	DIZER	1,323	0.06	230	99.14
33	TEMOS	1,165	0.05	231	99.57
34	ANOS	1,134	0.05	222	95.69
35	VER	1,132	0.05	229	98.71
36	ESTÃO	1,091	0.05	231	99.57
37	NOITE	1,088	0.05	218	93.97
38	VÍTIMA	1,082	0.05	204	87.93
39	BOM	1,072	0.05	223	96.12
40	SOU	1,040	0.05	227	97.84
41	CORPO	1,030	0.05	219	94.40

Figure 3b. Estatísticas sobre a série *CSI*.

N	Overall	C.S.I. - 1x01	1. - 1x02 - Cool.Chang	C.S.I. - 1x03 - Crate'n	Burial.srfs
text file					
file size	10,705,610	51,533	43,786		46,884
tokens (running words) in text	2,167,191	10,459	8,888		9,730
tokens used for word list	892,346	4,298	3,712		3,528
sum of entries					
types (distinct words)	34,627	1,352	1,154		1,148
type/token ratio (TTR)	3.88	31.46	31.09		32.54
standardised TTR	23.99	24.77	24.66		21.97
standardised TTR std.dev.	75.96	66.55	65.27		68.21
standardised TTR basis	1,000	1,000	1,000		1,000
mean word length (in characters)	3.18	3.16	3.16		3.04
word length std.dev.	1.93	1.89	1.87		1.76
sentences	147,650	800	727		730
mean (in words)	6.04	5.37	5.11		4.83
std.dev.	4.73	4.27	4.10		3.39
paragraphs	232	1	1		1
mean (in words)	3,846.32	4,298.00	3,712.00		3,528.00
std.dev.	608.17				
headings					
mean (in words)					
std.dev.					
sections	232	1	1		1
mean (in words)	3,846.32	4,298.00	3,712.00		3,528.00
std.dev.	608.17				
numbers removed	1,274,864	6,161	5,176		6,203
stoplist tokens removed	421,183				
stoplist types removed	4,222				
1-letter words	88,263	425	338		361
2-letter words	996,612	4,816	4,116		4,696
3-letter words	568,201	2,779	2,305		2,671
4-letter words	124,534	602	563		510
5-letter words	129,078	636	560		561
6-letter words	89,772	408	338		337
7-letter words	63,185	302	271		225
8-letter words	45,858	212	192		179
9-letter words	27,182	143	92		77
10-letter words	16,068	62	56		44
11-letter words	9,751	36	23		33
12-letter words	4,692	22	19		25
13-letter words	2,388	7	9		7

Figura 3a. Wordlist da série *CSI*.

Figura 3b. Estatísticas sobre a série *CSI*.

b. *Keywords*: através da ferramenta, pode-se elaborar uma análise contrastiva entre a listagem de palavras do *corpus* de estudo que foi compilado e a de um *corpus* de referência geral de língua<sup>5</sup> (no mínimo cinco vezes maior). Como resultado, temos os candidatos a termos do *corpus* em análise. A figura 4 mostra os candidatos a termos em língua inglesa para a série *Farscape*.

<sup>4</sup> Foi elaborado, para todo o projeto, um filtro com as palavras gramaticais do português e do inglês, já que as mesmas não aparecem como termos (quase sempre substantivos).

<sup>5</sup> Nesse caso, um amálgama entre o BNC (*British National Corpus*) e o ANC (*American National Corpus*).

N	Key word	Freq	%RC	RC	RC. %	Keyness	Lemmas	Sel
1	CRICHTON	1,223	0.14	57		11,713.44	0.0000000000	
2	AERYN	976	0.12	9		9,616.55	0.0000000000	
3	D'ARGO	830	0.10	2		8,236.78	0.0000000000	
4	MOYA	847	0.10	27		8,193.61	0.0000000000	
5	RYGEL	574	0.07	14		5,583.41	0.0000000000	
6	PILOT	1,069	0.13	3,837		5,554.24	0.0000000000	
7	TALYN	547	0.06	0		5,446.67	0.0000000000	
8	ZHAAN	519	0.06	5		5,111.44	0.0000000000	
9	CHIANA	484	0.06	5		4,763.61	0.0000000000	
10	SCORPIUS	438	0.05	14		4,236.60	0.0000000000	
11	CRAIS	341	0.04	0		3,395.37	0.0000000000	
12	PEACEKEEPER	327	0.04	9		3,173.18	0.0000000000	
13	SHIP	729	0.09	4,743		3,029.36	0.0000000000	
14	WORMHOLE	312	0.04	10		3,017.62	0.0000000000	
15	MOYA'S	237	0.03	2		2,336.72	0.0000000000	
16	PEACEKEEPERS	263	0.03	85		2,232.94	0.0000000000	
17	STARBURST	206	0.02	19		1,921.12	0.0000000000	
18	FRELL	215	0.03	46		1,898.31	0.0000000000	
19	SCARRAN	178	0.02	0		1,772.33	0.0000000000	
20	LUXAN	170	0.02	4		1,654.64	0.0000000000	
21	SCARRANS	164	0.02	0		1,632.93	0.0000000000	
22	HEY	446	0.05	3,834		1,632.90	0.0000000000	
23	LEVIATHAN	174	0.02	63		1,458.90	0.0000000000	
24	KILL	458	0.05	6,653		1,254.30	0.0000000000	
25	PLANET	303	0.04	2,090		1,227.61	0.0000000000	
26	STARK	233	0.03	950		1,159.20	0.0000000000	
27	JOHN	934	0.11	35,900	0.03	1,087.74	0.0000000000	
28	POD	167	0.02	293		1,064.11	0.0000000000	
29	FRELLING	131	0.02	65		1,056.20	0.0000000000	
30	MICROTS	103	0.01	0		1,025.55	0.0000000000	
31	WORMHOLES	107	0.01	5		1,024.59	0.0000000000	
32	GONNA	597	0.07	16,252	0.01	1,008.30	0.0000000000	
33	CRICHTON'S	109	0.01	13		1,002.70	0.0000000000	
34	ARNS	104	0.01	5		994.99	0.0000000000	
35	ARN	101	0.01	12		929.31	0.0000000000	
36	SEBACEAN	90	0.01	0		896.11	0.0000000000	
37	JOOL	89	0.01	0		886.16	0.0000000000	
38	DRDS	84		0		836.37	0.0000000000	
39	DIE	362	0.04	6,882		824.63	0.0000000000	
40	YAH	107	0.01	80		811.16	0.0000000000	
41	TALYN'S	81		0		806.50	0.0000000000	

Figura 4. Candidatos a termos em língua inglesa para a série *Farscape*.

c. *Concordance*: ao selecionarmos um candidato a termo na ferramenta *Keyword* e pedirmos suas linhas de concordância, uma listagem, em formato *KWIC* (*Key Word in Context*) nos é disponibilizada. Através dessas linhas, com a palavra chave centralizada em azul, podemos coletar todas as características do termo, das morfosintáticas às semânticas, além de subsídios para a criação da definição do mesmo. Abaixo, na figura 5, vemos as linhas de concordância para o termo *Leviathan*, da série *Farscape*. Para encontrarmos contextos definitórios (AUBERT, 1996), podemos usar o verbo *ser* (ou *to be*, em inglês). Na figura 6a, um exemplo para o termo *transporte* da série *Star Trek* (tomando todas as séries derivadas como *corpus*); na figura 6b, temos o arquivo da primeira linha de concordância, com o termo em evidência.

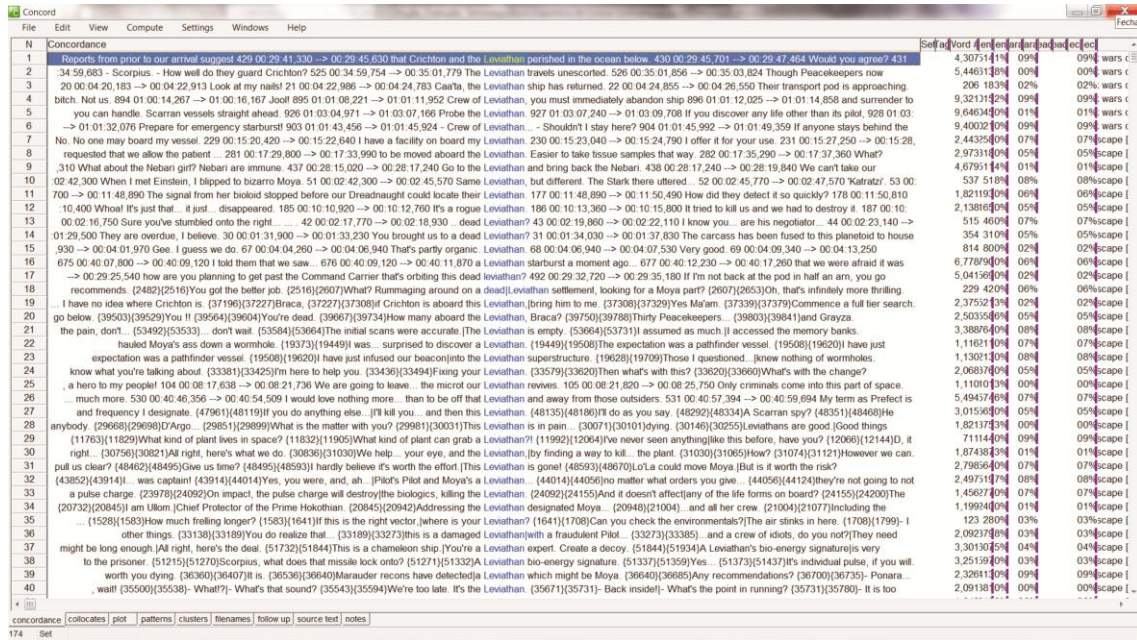


Figura 5. Termo *Leviathan*, série *Farscape*.

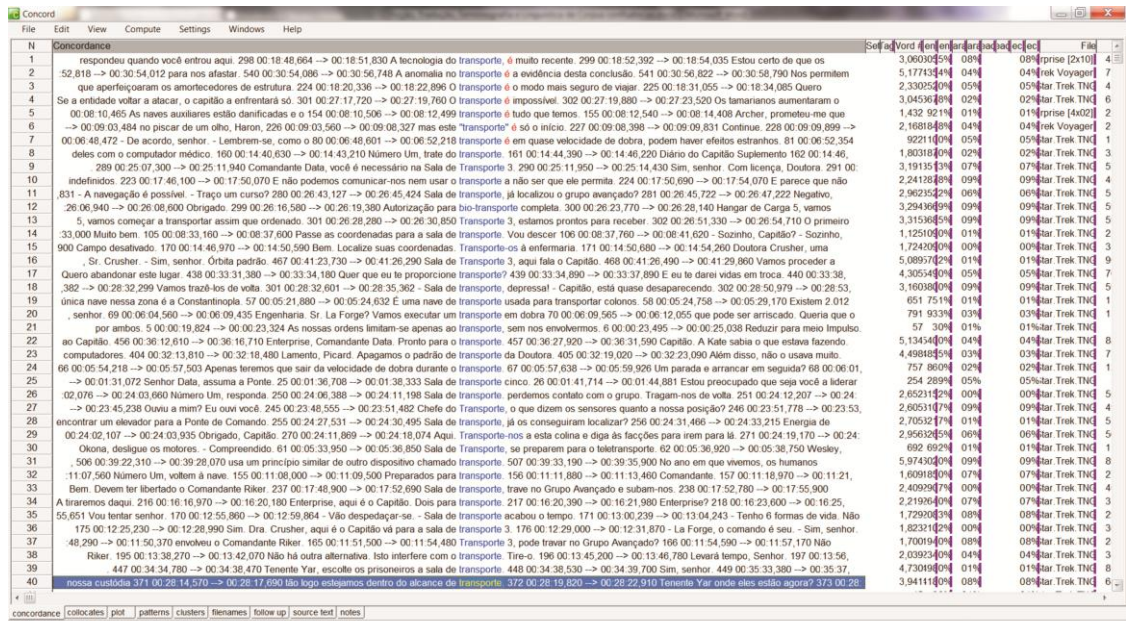


Figura 6a. Termo *transporte*, séries de *Star Trek*.

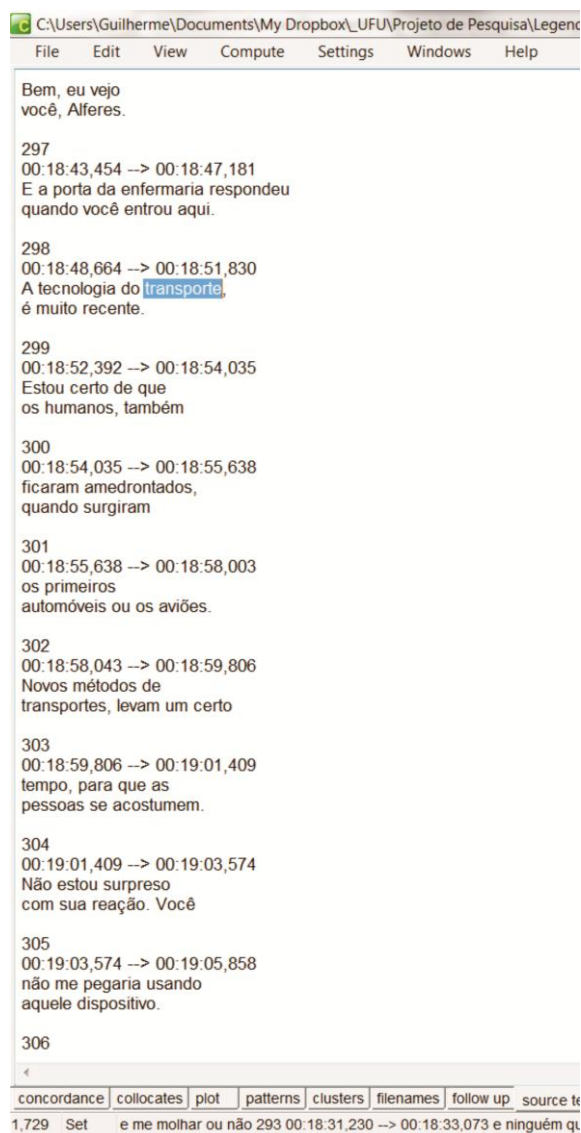


Figura 6b. Termo *transporte* (*Star Trek*) em destaque no texto da primeira linha de concordância da figura anterior.

## 5. A inserção dos termos no banco de dados

Terminado o levantamento das linhas de concordância de um termo e verificado se esse tem as condições de oferecer uma definição nas duas línguas em questão (inglês e português), o próximo passo é o cadastramento do mesmo no banco de dados do VoTec (página com acesso restrito). Os campos são preenchidos apenas se o termo fornecer as informações necessárias, caso contrário, serão deixados em branco.

Para esse projeto em específico, foi criado um novo site, dedicado apenas aos trabalhos dos alunos de Iniciação Científica. Na figura 7, podemos verificar a tela inicial do administrador do sistema. Aqui, aparecem os termos que estão sendo trabalhados pelos alunos. Assim que os mesmos são finalizados, o administrador os aprova e eles são disponibilizados na página de visualização.



Vocabulário Técnico Online Tela Cheia | English

Bem Vindo, Guilherme Fromm

Seu último acesso foi em: 02/11/2011 03:50

[Trocar Senha](#)
[Cadastro de Pesquisadores](#)
[Novo Termo](#)
[Cadastro de Áreas](#)
[Cadastro de Fontes](#)
[Sair](#)

### Fichas ainda não aprovadas

Termo	Língua	Coletado em	Cadastrada por	Ações
Amyloidosis		22/10/2011	Mila Bang	editar - revisar - reprovar - aprovar
Amiloidose		22/10/2011	Mila Bang	editar - revisar - reprovar - aprovar
Clot		22/10/2011	Mila Bang	editar - revisar - reprovar - aprovar
Coágulo		22/10/2011	Mila Bang	editar - revisar - reprovar - aprovar
Seizure		02/10/2011	Mila Bang	editar - revisar - reprovar - aprovar
Convulsão		02/10/2011	Mila Bang	editar - revisar - reprovar - aprovar
demônio		01/10/2011	Raphael Marco Oliveira Carneiro	editar - revisar - reprovar - aprovar
demon		01/10/2011	Raphael Marco Oliveira Carneiro	editar - revisar - reprovar - aprovar
Delvian		26/09/2011	Flávia Santos	editar - revisar - reprovar - aprovar
Frell		26/09/2011	Flávia Santos	editar - revisar - reprovar - aprovar
Antibiotic		25/09/2011	Mila Bang	editar - revisar - reprovar - aprovar
Antibiótico		25/09/2011	Mila Bang	editar - revisar - reprovar - aprovar
Wendigo		10/09/2011	Raphael Marco Oliveira Carneiro	editar - revisar - reprovar - aprovar
Scan		03/09/2011	Mila Bang	editar - revisar - reprovar - aprovar
Tomografia		03/09/2011	Mila Bang	editar - revisar - reprovar - aprovar
Prints		10/08/2011	Laura	editar - revisar - reprovar - aprovar
uma Mulher de Branco		04/08/2011	Raphael Marco Oliveira Carneiro	editar - revisar - reprovar - aprovar
a Woman in White		04/08/2011	Raphael Marco Oliveira Carneiro	editar - revisar - reprovar - aprovar
Wendigo		04/08/2011	Raphael Marco Oliveira Carneiro	editar - revisar - reprovar - aprovar

Total de Termos: 19

### Fichas aguardando revisão

Termo	Língua	Coletado em	Cadastrada por	Ações
Nenhum termo para revisar				

Total de Termos: 0

### Buscar fichas aprovadas

Termo:

02/11/2011 04:39 © 2007 FFLCH - ICMC Jr.

Figura 7. Termos a serem trabalhados pelos alunos.

Os passos a serem seguidos para a inserção dos termos são os seguintes:

a. Inserção das linhas de concordância escolhidas. No caso da figura 8, podemos visualizar, numa primeira tela do banco, as linhas de concordância escolhidas para compor o termo *Amyloidosis* (série *House*).

Vocabulário Técnico Online Tela Cheia | English

## Cadastro de Contextos para Amyloidosis

[Voltar ao Painel](#)   [Próximo Passo](#)

### Novo Contexto

#### Passo 2

Dados do Contexto

Exemplo\*:

Conceito\*:

Fonte\*: 5 Start Support Cadastrar

[Nova Fonte](#)

#### Contextos Cadastrados

Exemplo	Conceito	Fonte	Ações
"Yes. Can I go test that theory now? / Waste of time. He needs a bone marrow transplant for the amyloidosis."	It can be treated with a bone marrow transplant.	LTV 22/10/2011	<a href="#">editar</a> - <a href="#">excluir</a>
"Low immunoglobulin plus failing heart, plus fluid in the lungs, plus internal bleeding, equals amyloidosis."	It causes symptoms such as failing heart, low immunoglobulin, fluid in the lungs and internal bleeding.	LTV 22/10/2011	<a href="#">editar</a> - <a href="#">excluir</a>
"Kidneys are shutting down. Led to a sodium deficiency. Caused him to seize. Kidney failure means I was right about amyloidosis."	It is a disease that can cause kidney failure.	LTV 22/10/2011	<a href="#">editar</a> - <a href="#">excluir</a>
"Amyloidosis can cause seizures, and the protein buildup could cause organ failure."	It is a disease that can cause seizures.	LTV 22/10/2011	<a href="#">editar</a> - <a href="#">excluir</a>
"The problem was not the surgery. Abnormally deposited proteins. It is amyloidosis. That would explain the blinking and the low heart rate."	It is a disease that drops the heart rate and may cause constant blinking.	LTV 22/10/2011	<a href="#">editar</a> - <a href="#">excluir</a>
"Could be amyloidosis. Amyloid buildup hits the organs, tissue... her muscles would not have atrophied."	It is a disease which does not cause the muscles to atrophy.	LTV 22/10/2011	<a href="#">editar</a> - <a href="#">excluir</a>

Contextos Cadastrados: 6

02/11/2011 05:08 © 2007 FFLCH - ICMC Jr.

Figura 8. Contextos do termo *Amyloidosis* (série *House*).

b. A partir dos contextos cadastrados, o aluno vai para uma segunda página do banco e passa a preencher os campos disponíveis nas abas visíveis na parte inferior do quadro (figura 9, mesmo termo em português).

Vocabulário Técnico Online Tela Cheia | English

## Termo: Amiloidose

Contextos Conceito Final / Definição

Contextos

Exemplo	Conceito	Fonte
1 "Li todos os e-mails. Amiloidose foi a mais votada. / Se tivesse amiloidose, haveria manchas no seu ultrassom cardíaco."	É uma doença que apresenta manchas quando feito um exame de ultrassom.	LTV 22/10/2011
2 "Ele está alucinando porque está queimando. Amiloidose não faria isso."	É uma doença que não causa alucinações.	LTV 22/10/2011
3 "Não é um ganglioma. O problema não foi a cirurgia. Acúmulo anormal de proteínas. É amiloidose. Explicaria o piscar e a bradicardia."	É uma doença que pode causar bradicardia e piscar de olhos constante.	LTV 22/10/2011

Dados Traços Distintivos Semântica Termo Equivalente Termos Remissivos Informações Enciclopédicas

Dados

Ontologia: Terminologia em Ficção > House

Categoria Gramatical: 
   
 Número:

Gênero: 
   
 Sigla/Acrônimo:

Entrada por Extenso:

Var. Morfossintáticas:

Acepção N°:

02/11/2011 05:13 © 2007 FFLCH - ICMC Jr.

Figura 9. Termo *Amiloidose*, série *House*.

A aba Dados (figura 10) deve ser preenchida com as características morfossintáticas e a posição do termo no *corpus* de estudo.

Dados Traços Distintivos Semântica Termo Equivalente Termos Remissivos Informações Enciclopédicas

Dados

Ontologia: Terminologia em Ficção > House

Categoria Gramatical: 
   
 Número:

Gênero: 
   
 Sigla/Acrônimo:

Entrada por Extenso:

Var. Morfossintáticas:

Acepção N°:

Cópus

Posição na Ordem de Freqüência: 
   
 N° de Ocorrências do termo:

02/11/2011 05:13 © 2007 FFLCH - ICMC Jr.

Figura 10. Aba Dados, termo *Amiloidose*, série *House*.

A aba Traços Distintivos (figura 11) permite, através de uma Análise Componential (ILARI, 2003) a decomposição dos conceitos previamente elaborados pelos alunos em traços semânticos para ajudar na criação de uma definição.

	A	B	C	D	E	F	G	H	I	J
1	doença	apresenta	manchas	ultrassom						
2	doença				causa	alucinaçõe...				
3	doença				causa	bradicardi...				
4	amiloidose					falência r...	sintoma			
5	doença				causar	baixa da i...				
6	doença							tratamento	transplant...	medula óss..
7	doença				causar	convulsões				

Figura 11. Aba Traços Distintivos, termo *Amiloidose*, série *House*.

Na aba Semântica (figura 12), o aluno, além de poder inserir dados do termo se o mesmo já foi dicionarizado (esses dados não são disponibilizados na página de consulta), também estabelece, através dos exemplos, as relações semânticas ali presentes.

Semântica

Termo Dicionarizado?  Definições Coincidentes?

Fonte: Escolha uma fonte

Definição Dicionarizada:

Hiperônimo de:

Hipônimo de:

Co-hipônimo de:

Sinônimo(s):  Antônimo(s):

Notas: Termo não dicionarizado. Encontrado no Wiktionary, no Urban Dictionary, The Free Dictionary.

02/11/2011 05:45 © 2007 FFLCH - ICMC Jr.

Figura 12. Aba Semântica, termo *Wendigo*, série *Supernatural*.

Na aba Termo Equivalente (figura 13), liga-se o termo em análise ao mesmo termo na outra língua.

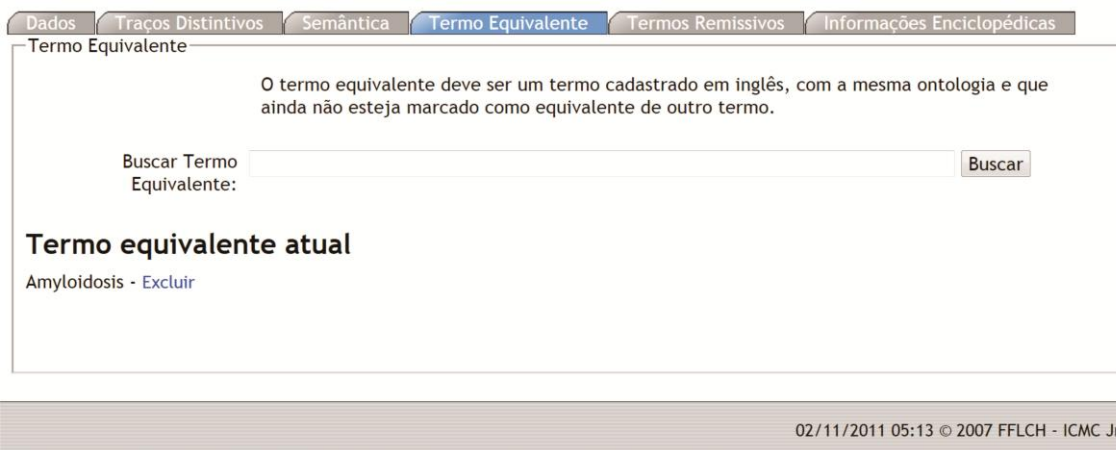


Figura 13. Aba Termo Equivalente, termo *Amiloidose*, série *House*.

A aba Termos Remissivos (figura 14) permite que o aluno amplie o entendimento do termo em análise, desde que os termos remissivos já estejam cadastrados no sistema e que apareçam nos exemplos do termo em questão.

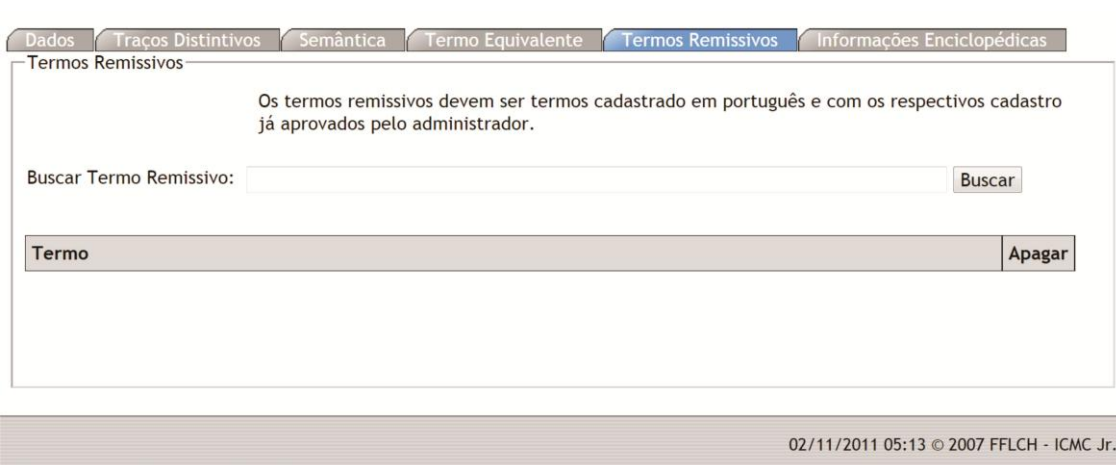


Figura 11. Aba Termos Remissivos.

As Informações Enciclopédicas (figura 15), normalmente, são retiradas da Wikipédia. Como essas informações são disponibilizadas no site do VoTec, não podem provir de outros sites com direito autoral.

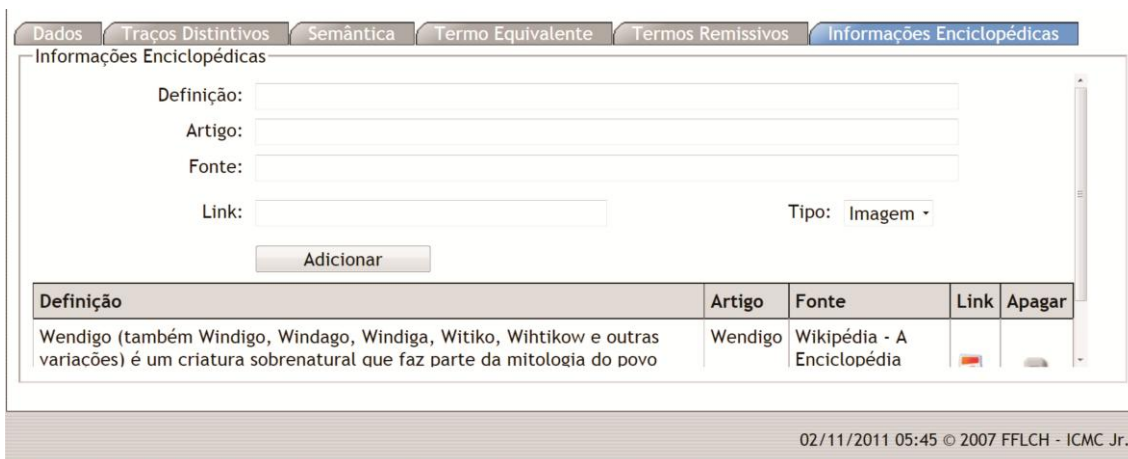


Figura 12. Informações Enciclopédicas, termo *Wendigo*, série *Supernatural*.

Após a inserção de todos esses termos, o aluno cria um conceito final (figura 16) para o seu termo e o transforma numa definição (esse processo pode ser uma cópia ou um retrabalho, visando facilitar a compreensão por parte do consulente).

The screenshot shows the 'Vocabulário Técnico Online' interface. At the top, it says 'Vocabulário Técnico Online' on the left and 'Tela Cheia | English' on the right. Below this is a header bar with the term 'Termo: Wendigo'. Underneath the header are three buttons: 'Passo Anterior', 'Salvar', and 'Sair sem salvar'. Below the buttons are two tabs: 'Contextos' and 'Conceito Final / Definição', with the latter being selected. The main content area is titled 'Conceito Final / Definição' and contains two text boxes. The first box is labeled 'Conceito Final:' and contains the text: 'Wendigo é um monstro que hiberna durante anos mantendo suas vítimas vivas e estocadas para comê-las depois. Wendigos têm garras e imitam a voz humana. Um Wendigo, antes de ser um mostro, foi um homen que, para sobreviver durante um inverno rigoroso, devorava os membros de sua tribo ou grupo. Um Wendigo não pode ultrapassar Símbolos Anasazi.' The second box is labeled 'Definição:' and contains the same text: 'Wendigo é um monstro que hiberna durante anos mantendo suas vítimas vivas e estocadas para comê-las depois. Wendigos têm garras e imitam a voz humana. Um Wendigo, antes de ser um mostro, foi um homen que, para sobreviver durante um inverno rigoroso, devorava os membros de sua tribo ou grupo. Um Wendigo não pode ultrapassar Símbolos Anasazi.'

Figura 16. Conceito Final/Definição, termo *Wendigo*, série *Supernatural*.

## 6. A visualização na página de consulta

Terminados todos os passos anteriores, o termo é disponibilizado na página de consulta. Importante notar, aqui, que os termos não necessariamente são uma tradução (embora o *corpus*, nesse caso, seja paralelo). O site nos permite visualizar o termo em contraste nas duas línguas, mas a construção dos mesmos em cada língua é independente, dependendo dos exemplos coletados. Essa informação fica mais clara se consultados os termos no outro site do VoTec, dedicado a projetos diversos (<http://www.pos.voteconline.com.br/>). Na figura 17, podemos visualizar o termo *Wendigo*, da série *Supernatural*.

Vocabulário Técnico Online Tela Cheia | English | Ajuda

Terminologia em Ficção ▾ Escolha uma área ▾

**Tipos de Exibição**

Normal  
Descritiva

---

**Tipos de Consulta**

Total  
Tradutor  
Modular

---

**Consultas Externas**

Corpus NILC  
Google  
Answers.com  
Wikipedia  
CORTEC

▼ Português

[Voltar ao resultado da busca](#)

**Wendigo. Sobrenatural. s.m.s.** Wendigo é um monstro que hiberna durante anos mantendo suas vítimas vivas e estocadas para comê-las depois. Wendigos têm garras e imitam a voz humana. Um Wendigo, antes de ser um monstro, foi um homem que, para sobreviver durante um inverno rigoroso, devorava os membros de sua tribo ou grupo. Um Wendigo não pode ultrapassar Símbolos Anasazi.. Ex.: Wendigo é um termo nativo americano. Significa "o mal com poderes". Tem mais de cem anos. Todos já foram homens. Algumas vezes um índio, ou um minerador ou caçador. - Como alguém se transforma nisso? - É sempre a mesma coisa. Durante um inverno rigoroso, o homem começa a passar fome sem poder pedir por ajuda. Então ele vira um canibal para sobreviver. Come outros membros do seu grupo.. *Corpus: Posição na Ordem de Frequência: (4.256); Nº de Ocorrências do termo: (6).* Informações Enciclopédicas: Wendigo (também Windigo, Windago, Windiga, Witiko, Wihtikow e outras variações) é um criatura sobrenatural que faz parte da mitologia do povo indígena da América do Norte Ojibwas. De acordo com a mitologia, o Wendigo é formado a partir de um humano qualquer Em: [Wendigo - Wikipédia - A Enciclopédia Livre](#)

▼ English

[Go back to search results](#)

**Wendigo. Supernatural. n.m/f.s.** Wendigo is a monster that hibernates for years keeping its victims alive and storing them to eat later. It usually lives in woods. It has claws and it can mimic the human voice. A Wendigo was once a man that in order to survive during harsh winters devoured the members of his tribe or camp. It cannot cross over Anasazi symbols.. Ex.: "Wendigo" is a Cree Indian word.It means "evil that devours." They're hundreds of years old. Each one was once a man. Sometimes an Indian, or other times a frontiersman or a miner or a hunter. -How's a man turn in to one of those? DEAN: Well, it's always the same. During some harsh winter, a guy finds himself starving...cut off from supplies or help. He becomes a cannibal to survive, eating other members of his tribe or camp.. *Corpus: Frequency order position: (2.977); Term number of occurrences: (8).* Encyclopedic Information: The Wendigo (also known as Windigo, Weendigo, Windago, Waindigo, Windiga, Witiko, Wihtikow, and numerous other variants)[1] is a mythical creature appearing in the mythology of the Algonquian people. It is a malevolent cannibalistic spirit which can transm: [Wendigo - Wikipedia, the free encyclopedia](#)

02/11/2011 04:24 © 2007 Guilherme Fromm - ICMC Jr.

Figura 17. Página de visualização, termo *Wendigo*, série *Supernatural*.

## 7. Considerações Finais

A Terminologia, longe de estar disponível apenas nas áreas de especialidade, já alcança o dia a dia de todos. Até nos momentos de relaxamento, quando assistimos televisão, elas se fazem presentes. O projeto Terminologia em Ficção, além de demonstrar essa realidade, serve como ponto de partida para os alunos de Iniciação Científica iniciarem seus treinamentos nas áreas de Linguística de *Corpus* (compilação e análise de *corpora*) e Terminografia Bilíngue (inglês/português).

Com a experiência adquirida no projeto, esses alunos já estão habilitados a fazer parte de projetos terminológicos diversos e/ou propor seus próprios projetos para um futuro mestrado na área de descrição linguística.

## 8. Referências Bibliográficas

AUBERT, F. H. **Introdução à metodologia da pesquisa terminológica bilíngüe**. São Paulo: Humanitas, 1996.

BANG, M. **Vocabulário de Especialidade na Ficção**: análise do uso de termos da área médica na Série *House M.D.* Iniciação Científica, em curso. 2011. Orientador: Prof. Dr. Guilherme Fromm.

BARBOSA, M. A. Dicionário, vocabulário, glossário: concepções. In: ALVES, I. M. (org.). **A constituição da normalização terminológica no Brasil**. São Paulo: FFLCH/CITRAT, 2001.

BARROS, L.A.. Aspectos epistemológicos e perspectivas científicas da terminologia. **Cienc. Cult.** [online]. 2006, v. 58, n. 2, pp. 22-26.

BERBER-SARDINHA, A. **Linguística de Corpus**. São Paulo: Manole, 2004.

CARNEIRO, R. M. O. **Linguística de Corpus e Ficção**: uma análise contrastiva bilíngue do vocabulário de especialidade na série *Supernatural*. Iniciação Científica, em curso. 2011. Orientador: Prof. Dr. Guilherme Fromm.

FROMM, G. **VoTec**: a construção de vocabulários eletrônicos para aprendizes de tradução. São Paulo, 2007. Tese (Doutorado em Estudos Linguísticos e Literários em Língua Inglesa). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo.

ILARI, R. **Introdução ao estudo do léxico**: brincando com as palavras. 2<sup>a</sup>. ed. São Paulo: Contexto, 2003.

LAGO, L. P. **Análise de Vocabulário de Especialidade na Ficção na Série CSI**: um estudo qualitativo de tradução. Iniciação Científica, em curso. 2011. Orientador: Prof. Dr. Guilherme Fromm.

MARTHE, M. A cura pela razão. **Revista Veja**, 16/03/2011.

SAYURI, J. Legendários. **Revista Superinteressante**, maio 2011.

SILVA, F. S. **Uma viagem ao corpus de Farscape**: as questões de tradução envolvidas em um *corpus* trilíngue de ficção científica. Iniciação Científica, em curso. 2011. Orientador: Prof. Dr. Guilherme Fromm.

SCOTT, M. **WordSmith Tools version 5**. Liverpool: Lexical Analysis Software, 2008.