

***Ferramentas  
Lingüísticas  
para o BrOffice***

# ***Verificador Ortográfico***

- ◆ O Verificador Ortográfico do BrOffice alcança a marca de 1.400.000 palavras, e estará disponível para download já no dia do evento.
- ◆ O projeto conta, atualmente, com **40** colaboradores, sendo que Raimundo Santos Moura é o coordenador.

# ***Verificador Ortográfico***

<http://www.openoffice.org.br/?q=verortografico>

**raimundomoura@openoffice.org**

# ***DicSin***

## ***Dicionário de Sinônimos***

- ◆ BrOffice.org tem agora ferramenta própria para gerenciamento de Dicionário de Sinônimos.
- ◆ O **DicSin** permite o controle do Dicionário de Sinônimos via Web.
- ◆ Incluir novas palavras / sinônimos e fazer correções.

# ***DicSin***

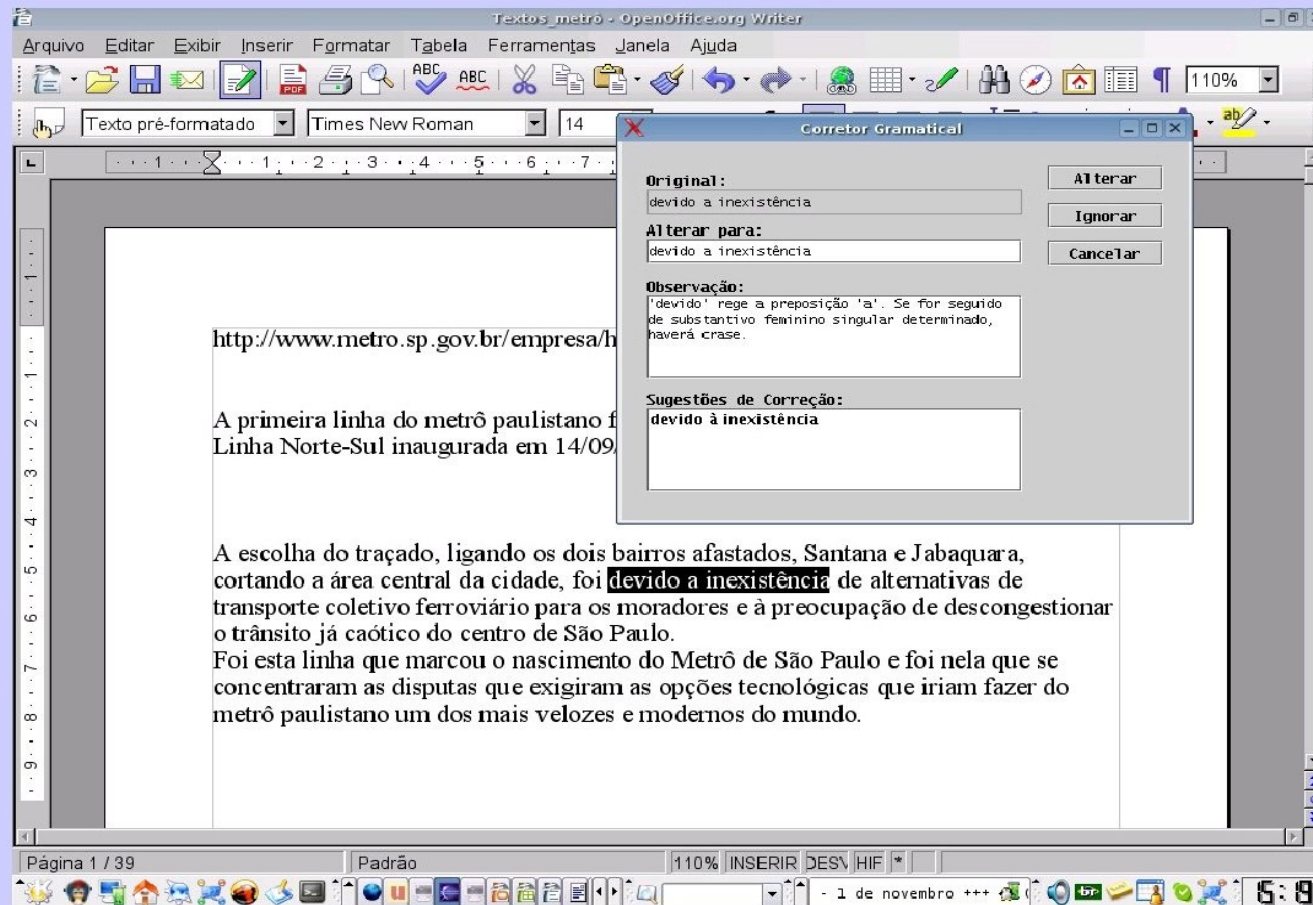
## ***Dicionário de Sinônimos***

### ◆ Colaboradores

- ◆ Fabio Souza
- ◆ Murillo Lamas
- ◆ Edgar Nogueira de Oliveira

<http://www.openoffice.org.br/?q=node/38>

# CoGrOO – Corretor Gramatical para a língua portuguesa, acoplável ao OpenOffice/BrOffice



# ***Introdução***

- ◆ Importância da ferramenta OO
  - ◆ Quantidade de downloads
  - ◆ Recursos / Usabilidade
  - ◆ Formatos livres
- ◆ Necessidade de um corretor gramatical para a língua Portuguesa
  - ◆ Devido à migração em massa ocorrendo nos órgãos públicos, faz-se necessário tal ferramenta.

# ***Introdução / História***

- ◆ Alternativas
  - ◆ Utilizar software proprietário (problemas já amplamente conhecidos/discutidos). Ex.:  
Freedows
  - ◆ Apoiar/desenvolver um software livre
- ◆ Projeto patrocinado pela FINEP:  
Chamada Pública MCT / FINEP /  
Software Livre
- ◆ Optamos pela última!!

# ***Objetivos***

- ◆ Obter um servidor de correção gramatical para pt-BR, acoplável à ferramentas de desktop
- ◆ Menor quantidade possível de falsos positivos
- ◆ Lançamento em maio/2006
- ◆ Proposta quanto ao acerto na correção: 30% de diferença com relação ao corretor do Microsoft

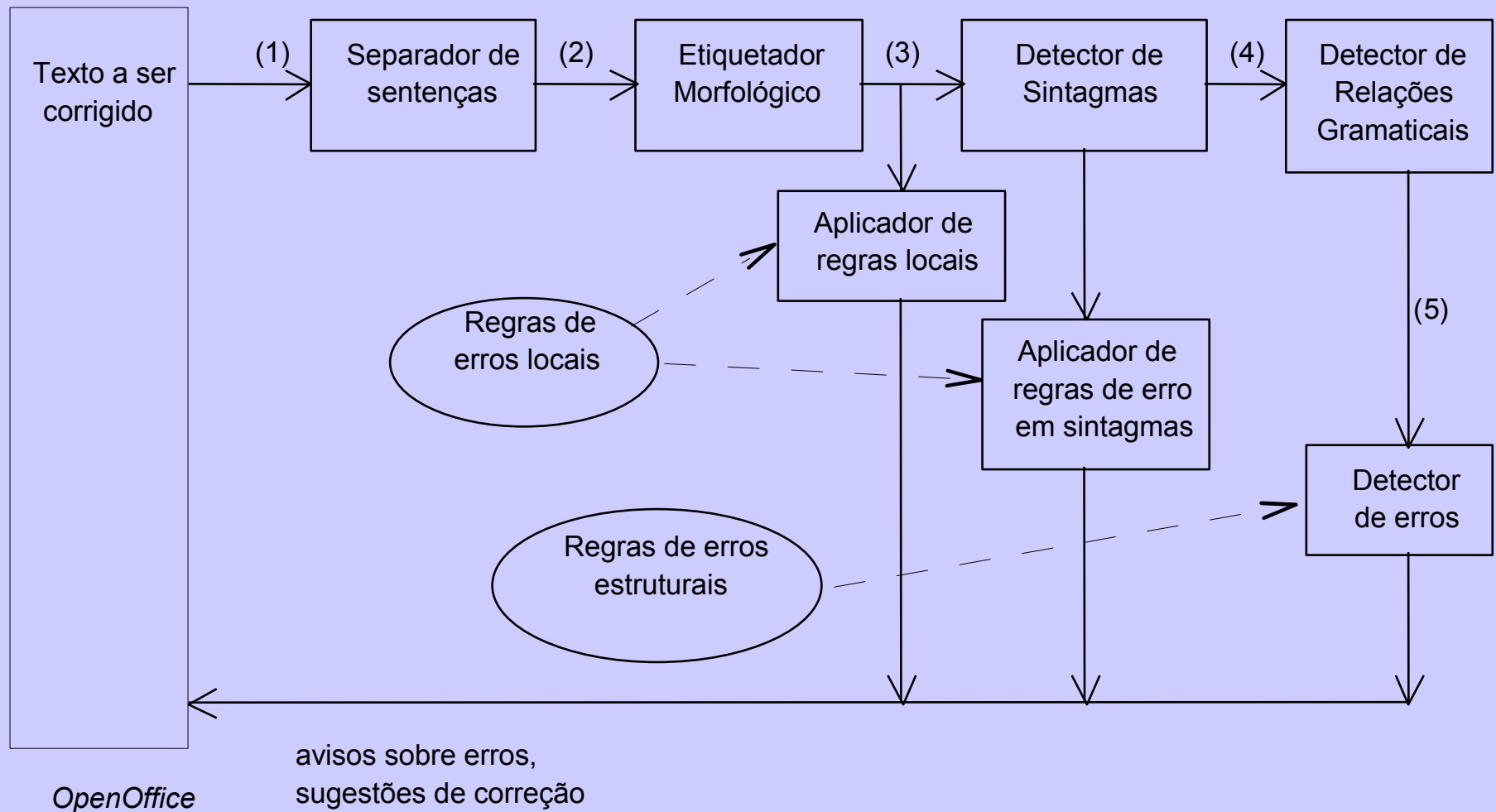
# ***Objetivos***

- ◆ Aplicação do CoGrOO na detecção de erros gramaticais:
  - ◆ concordância nominal e verbal;
  - ◆ crase (uso combinado de preposição e artigo);
  - ◆ regência nominal e verbal;
  - ◆ outros erros comuns na língua portuguesa falada no Brasil.

# ***Vantagens Estratégicas***

- ◆ **Abordagem híbrida:**
  - ◆ **Processamento estatístico de corpus lingüísticos**
  - ◆ **Regras sobre erros**
- ◆ **Inovação:**
  - ◆ **Primeiro corretor gramatical em software livre para a língua portuguesa**
  - ◆ **Dos mais robustos dentre os existentes**

# Arquitetura



# ***Separador de Sentenças***

**Entrada (1):    Ele foi procurar uma casa. Ele vai se casar com a Sra. Maria.**

**Saída (2):      Ele foi procurar uma casa . (1ª. sentença)  
Ele vai se casar com a Sra. Maria . (2ª. sentença)**

# ***Etiquetador Morfológico***

<b>Palavras</b>	<b>Etiquetas Morfológicas</b>
Ele	PERS_M_3S_NOM_(16740) PERS_M_3S_NOM/PIV_(62) N_M_S_(6)
foi	{ir}_V_PS_3S_IND_VFIN_(2753) {ser}_V_PS_3S_IND_VFIN_(55500) ADV_(138)
procurar	{procurar}_V_FUT_1/3S_SUBJ_VFIN_(1) {procurar}_V_INF_(759) {procurar}_V_INF_0/1/3S_(10) {procurar}_V_INF_3S_(33) {procurar} V_FUT_3S_SUBJ_VFIN_(15)
uma	NUM_F_S_(12037) DET_F_S_(96697) NUM_M/F_P_(32) NUM_M_P_(2) NUM_F_P_(16)
casa	{casar}_V_PR_3S_IND_VFIN_(114) N_F_S_(6034) {casar}_V_IMP_2S_VFIN_(1)
\$.	-PNT_NS(1108953)

# *Aplicação de regras locais*

## **"à" N\_M\_S\_**

**mensagem:** O sinal indicativo de crase indica que temos “a + a” expressos em um só “à”. Somente ocorre crase quando há encontro de preposição a com artigo ou pronome demonstrativo a/as.

**adequado:** Refiro-me ao trabalho remunerado./Refiro-me a trabalho remunerado.

**inadequado:** Refiro-me à trabalho remunerado.

## **"(meia|meias)" \_ADJ**

**mensagem:** A palavra “meio” usada no sentido de “um pouco” é advérbio, portanto invariável. Exceção: A palavra “meia” usada como substantivo concorda com o adjetivo.

**adequado:** A conclusão está meio confusa/ As conclusões estão meio confusas.

**inadequado:** A conclusão está meia confusa/ As conclusões estão meias confusas.

# ***Detector de sintagmas***

	<b>Palavras</b>	<b>Etiquetas Sintáticas</b>
Ele:[PERS_M_3S_NOM_]		
foi:[{ir}_V_PS_3S_IND_VFIN_]	Ele	NP_3S_M_0_
procurar:[{procurar}_V_INF_]	foi	VP_3S_0_
uma:[DET_F_S_]	procurar	
casa:[N_F_S_]	uma	NP_3S_F_1_
\$.:[-PNT_NS]	casa	
	.	-PNT_NS

# *Aplicação de regras de erro em sintagmas*

## **N\_F\_S\_ ADJ\_F\_P\_**

**mensagem:** O adjetivo concorda em gênero (masculino ou feminino) e número (singular ou plural) com o substantivo a que se refere.

**adequado:** Faces rosadas da criança ou a face rosada da criança.

**inadequado:** Face rosadas da criança.

## **DET\_M\_S\_ (N\_M\_P\_|N\_F\_S\_|N\_F\_P\_)**

**mensagem:** Os artigos definidos (o, a, os, as) concordam em número (singular ou plural) e em gênero (masculino ou feminino) com o substantivo a que se referem.

**adequado:** Os copos caíram da mesa. A roupa está no guarda-roupa. As roupas estão no guarda-roupa.

**inadequado:** O copos caiu da mesa. O roupa está no guarda-roupa. O roupas estão no guarda-roupa.

# ***Detector de relações gramaticais***

- ◆ Relações gramaticais: sujeito-verbo
- ◆ Busca de padrões no texto anotado, por exemplo: “! NP VP”
- ◆ Atribuição das etiquetas sintáticas:
  - ◆ SUBJ> associada a NP
  - ◆ (M)VP associada a VP

# ***Detector de relações gramaticais***

Ele foi procurar uma casa.

```
| - [ [SUBJ>]_M_3S_ ]
  | - [ (SJ>)SN_M_3S_0_ ]
    | - [ SN_M_3S_0_ ]
      | - PERS_M_3S_NOM_ --> Ele
| - [ (M)VP_3S_0_ ]
  | - {ser}_V_PS_3S_IND_VFIN_ --> foi
  | - {procurar}_V_INF_ --> procurar
| - [ SN_F_3S_1_ ]
  | - DET_F_S_ --> uma
  | - N_F_S_ --> casa
| - -PNT_ABS --> $.
```

# *Processo de Correção Gramatical*

Sentença: Os corretores foi procurar uma casa.

```
| - [ [SUBJ>]_M_3P_ ]
  | - [ (SJ>)SN_M_3P_0_ ]
    | - [ SN_M_3P_0_ ]
      | - DET_M_P_ --> Os
      | - N_M_P_ --> corretores
| - [ (M)VP_3S_0_ ]
  | - {ser}_V_PS_3S_IND_VFIN_ --> foi
  | - {procurar}_V_INF_ --> procurar
| - [ SN_F_3S_1_ ]
  | - DET_F_S_ --> uma
  | - N_F_S_ --> casa
| - -PNT_ABS --> $.
```

Mensagem de erro: Verificou-se erro de concordância entre o verbo e o sujeito.

# Resultados

<b>ERROS GRAMATICAIIS</b>	<b>CoGr00</b>	<b>ReGra</b>
Verdadeiros Positivos	14	15
Falsos Positivos	10	36

*Tabela 1: Erros Gramaticais detectados pelo CoGr00 e pelo ReGra*

<b>ERROS ESTILÍSTICOS</b>	<b>CoGr00</b>	<b>ReGra</b>
Verdadeiros Positivos	-	10
Falsos Positivos	-	2

*Tabela 2: Erros Estilísticos detectados somente pelo ReGra*

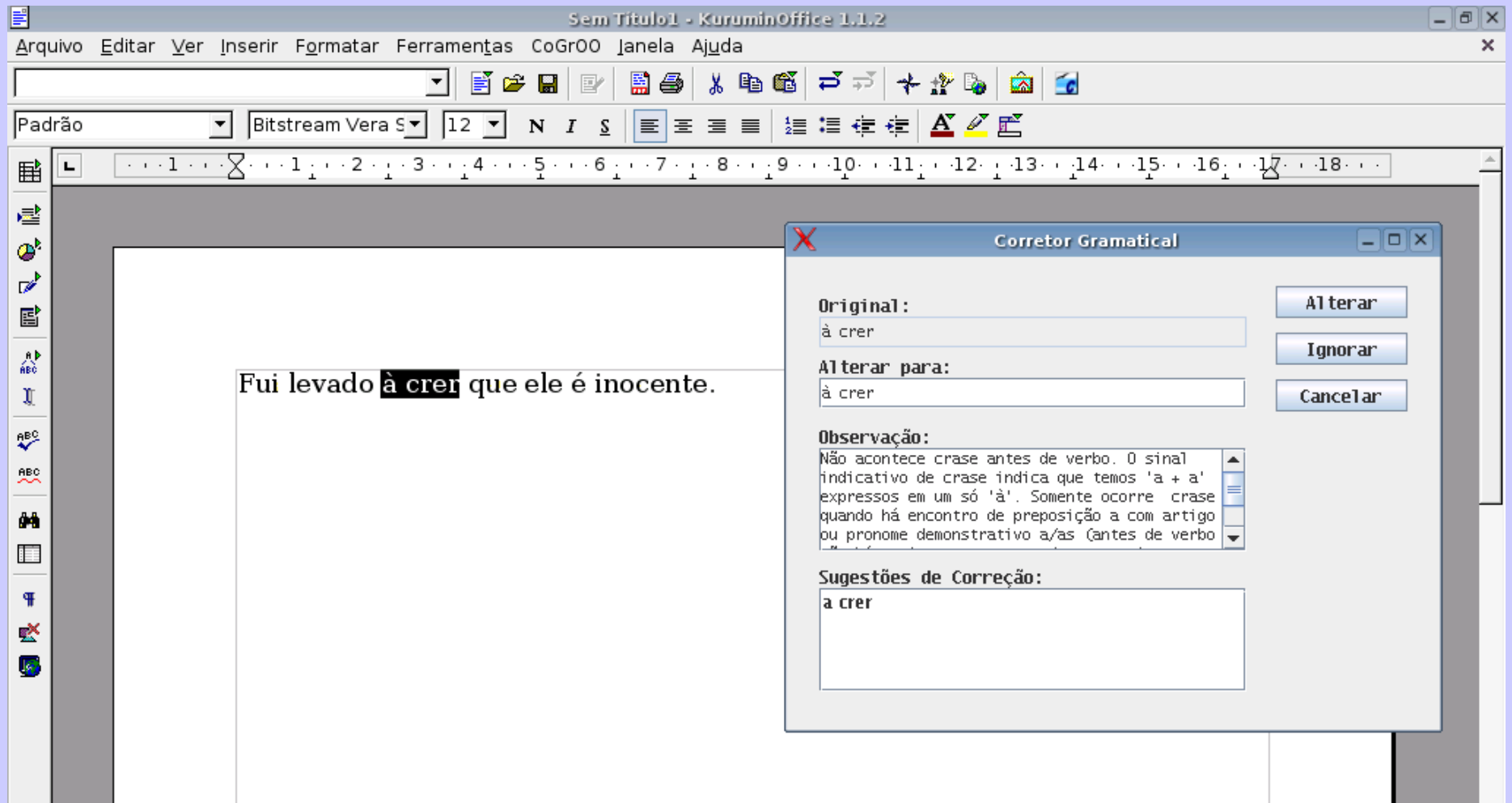
# ***Casos Contemplados***

- ◆ **crase (ignorando a análise da regência dos verbos);**
- ◆ **casos de concordância nominal (adjetivo com substantivo na ordem direta, tanto na função de adjunto adnominal como de predicativo de sujeito);**
- ◆ **emprego da expressão invariável "em anexo";**
- ◆ **emprego de "meio" na função de advérbio;**
- ◆ **concordância em gênero entre adjetivo e substantivo no adjunto adnominal;**

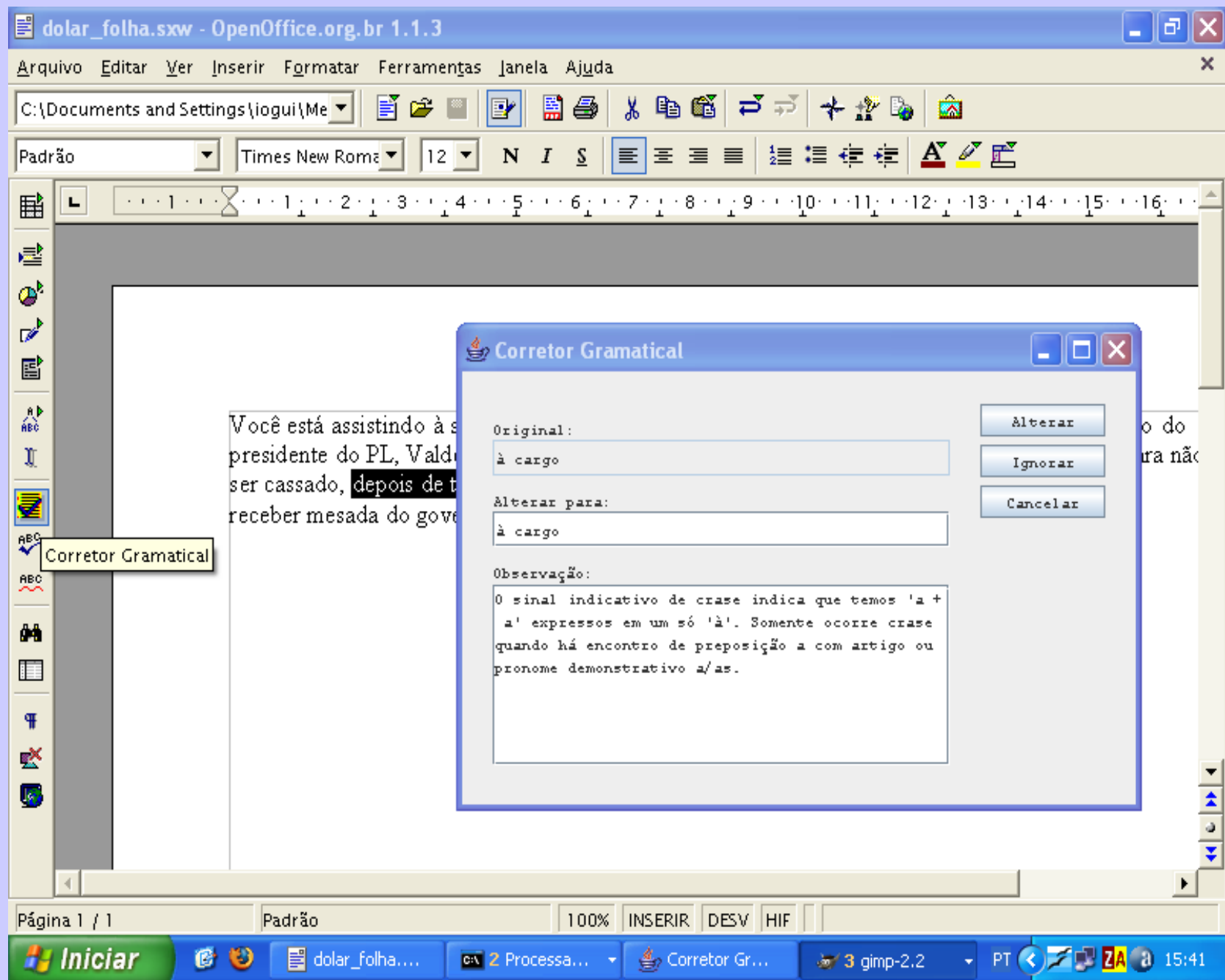
# ***Casos Contemplados***

- ◆ **flexão do verbo “fazer” indicando tempo;**
- ◆ **flexão do verbo “haver” indicando tempo e indicando existência;**
- ◆ **concordância de artigo com substantivo;**
- ◆ **uso dos pronomes “mim” e “ti”;**
- ◆ **uso de “mal” e “mau”;**
- ◆ **regência do verbo “preferir”.**

# Interface com OpenOffice



**OpenOffice 1.1.2 rodando no Linux**



## ***OpenOffice 1.1.3 rodando no Windows***

# ***Tratamento de Regras via Web***

- Atualizar regra
- Submeter nova regra
- Visualizar regras implementadas

Atualização de Regra - Mozilla Firefox

Arquivo Editar Exibir Ir Favoritos Ferramentas Ajuda

http://www.pcs.usp.br/~cogroo/rulesUpdate.cgi

Atualização de Regra Welcome to Gmail

# Atualização de Regra

## Regra 95

Método	local
Tipo	Concordância Artigo-Substantivo
Grupo	artigo plural + substantivo singular
Expressão Regular	DET_@_P_ N_@_S_
Mensagem	Os artigos definidos (o, a, os, as) e os artigos indefinidos (um, uma, uns, umas) concordam em número (singular ou plural) e em gênero (masculino ou feminino) com o substantivo a que se referem.
Marcar	lo o
Exemplos de sentenças	
Incorreta 1	Uns copo estão quebrados.
Correta 1	Uns copos estão quebrados.
Incorreta 2	Os copo estão quebrados.
Correta 2	Os copos estão quebrados.
Incorreta 3	As criança foram ao parque.
Correta 3	As crianças foram ao parque.
<input type="button" value="Remover exemplo"/> <input type="button" value="Adicionar exemplo"/>	
<input type="button" value="Limpar"/> <input type="button" value="Enviar"/>	

Concluído

# Submissão de Regra Nova

Método	Escolha método ▾
Tipo	Escolha tipo ▾
Grupo	Escolha grupo ▾
Expressão Regular	<input type="text"/>
Mensagem	<input type="text"/>
Marcar	<input type="text" value="0 0"/>
Exemplos de sentenças	
Incorreta 1	<input type="text"/>
Correta 1	<input type="text"/>
<input type="button" value="Adicionar exemplo"/>	
<input type="button" value="Limpar"/> <input type="button" value="Enviar"/>	

[Clique aqui para inserir regras](#)

[Clique aqui para atualizar regras](#)

[Clique aqui para checar se uma regra está funcionando](#)


[Perguntas mais frequentes \(FAQ\)](#)

Regras Implementadas — Corretor Gramatical para o OpenOffice - Mozilla Firefox

Arquivo Editar Exibir Ir Favoritos Ferramentas Ajuda

http://cogroo.incubadora.fapesp.br/portal/projeto/regrasImplementadas

Gmail - Screenshots Regras Implementadas — Corret...



# CoGrOO

projeto área de testes screenshots downloads equipe fórum suporte

você não está autenticado acessar cadastrar-se

você está aqui: página inicial » projeto » regras implementadas

context

- Imagens
- Regras Implementadas

notícias

- CoGrOO agora roda também no Windows! 24/08/2005
- CoGrOO no pré-FISL 6.0 07/06/2005

Mais...

## Regras Implementadas

Todas as regras atualmente disponíveis tem suas sentenças de exemplo submetidas ao CoGrOO e é possível detectar quais estão funcionando. A expressão regular e outras informações técnicas estão acessíveis ao se passar o ponteiro do mouse sobre o número da regra.

Grupo:	haver no sentido de existir
Mensagem:	O verbo haver no sentido de existir é impessoal. Permanece sempre na terceira pessoa do singular.
Regra	Exemplos de sentenças
50	Incorreta 1 <b>Haviam ratos</b> na cozinha.
	Correta 1 Havia ratos na cozinha.
	Incorreta 2 <b>Haviam nuvens</b> escuras no céu.
	Correta 2 Havia nuvens escuras no céu.
	Incorreta 3 <b>Haverão ratos</b> na cozinha.
	Correta 3 Haverá ratos na cozinha.
	Incorreta 4 <b>Haverão nuvens</b> escuras no céu.
	Correta 4 Haverá nuvens escuras no céu.
Grupo:	verbo auxiliar + haver no sentido de existir
Mensagem:	Verbo auxiliar + "haver" no sentido de existir é impessoal. Permanece sempre na terceira pessoa do singular.
Regra	Exemplos de sentenças

« Out 2005 »

D	S	T	Q	Q	S	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

Concluído

# ***Conclusões***

- ◆ Comparação com Regra, corretor gramatical da língua portuguesa usado no Microsoft Office, mostra que o projeto CoGrOO é promissor.

# ***Próximas Ações***

- ◆ Melhorar a interface com o usuário (Google SoC – Construção de API para Grammar Checkers)
- ◆ Melhorar o motor de correção
  - ◆ **Sugestões de correção**
  - ◆ **Regras para Estilo (quase pronto!)**
- ◆ CoGrOO 2.0 em fase de projeto
- ◆ Portar para outras línguas (inglês, espanhol)

# Créditos

## **Desenvolvedores:**

*William Colen*

*Marcos Oku*

*Marcelo Suzumura*

*Fábio Wang*

*Fábio Blanco*

*Bruno Sant'Anna*

*Sueli Uliano*

*Edgard Lemos Jr.*

*Lais Salvador*

*Carlos Menezes*

*Jorge Kinoshita*

## **Executores:**

**USP – SENAC**

## **Patrocinador:**

**FINEP**

## **Incubadora:**

**FAPESP**

## **Empresa apoiadora:**

**Metrô-SP**



# ***Perguntas?***

## ***Nosso portal:***

<http://cogroo.incubadora.fapesp.br>

## ***E-mail:***

[carlos.edmenezes@sp.senac.br](mailto:carlos.edmenezes@sp.senac.br)