

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**

**MÔNICA RIGO AYRES**

**AELIUS FALADO 1.0**

**PORTO ALEGRE**

**2014**

MÔNICA RIGO AYRES

**AELIUS FALADO 1.0**

Monografia apresentada ao Instituto de Letras do Rio Grande do Sul como requisito parcial para a conclusão do curso de Licenciatura em Letras.

Prof. Dr. Gabriel de Ávila Othero  
Orientador

PORTO ALEGRE  
2014

## AGRADECIMENTOS

Gostaria de agradecer carinhosamente:

Ao professor Gabriel de Ávila Othero, que me orientou nesta e em outras pesquisas, sempre me mostrando que caminho seguir e confiando na minha capacidade e independência.

Ao professor Leonel de Alencar e seu grupo de pesquisa, especialmente seu bolsista Hélio, da Universidade Federal do Ceará, pela parceria no projeto.

À professora Gisela Collischonn e toda equipe do Varsul, pelos trechos de língua falada gentilmente disponibilizados para minha pesquisa.

Aos demais professores do Instituto de Letras, que sempre contribuíram para meu crescimento.

Aos colegas do curso de Letras, que sempre discutiram ideias comigo e dividiram momentos de aprendizado, especialmente: Thiely, Ana Carolina, Victor, Marianna, Isabel e Camila.

Ao CNPq, pela concessão de bolsas de Iniciação Científica que financiaram meus estudos; à UFRGS, pela oportunidade de estudo.

À minha família, que apesar de longe fisicamente sempre esteve e está comigo em todos os momentos, me incentivando a ir além e alcançar meus objetivos.

Ao meu namorado, Fernando, pelo apoio e carinho.

A Deus, por ter me dado capacidade para concluir este trabalho, e por ter me dado a bênção e o privilégio de conhecer todas essas pessoas que me ajudaram.

## RESUMO

Este trabalho insere-se nas áreas de Linguística de Corpus e análise morfológica do português brasileiro, tratando do tema anotação automática morfológica de *corpus* de língua falada. O trabalho pretende, entre outras coisas, contribuir com a melhoria do etiquetador automático morfossintático Aelius, desenvolvido originalmente pelo prof. Dr. Leonel Alencar, da Universidade Federal do Ceará, coordenador do projeto CompLin – Computação e Linguagem Natural. As etiquetas utilizadas no Aelius são as mesmas do *corpus* de português histórico Tycho-Brahe (*corpus* eletrônico já anotado, composto de textos em português escritos por autores nascidos entre 1380 e 1845). Os textos que utilizamos para a anotação são do banco do projeto Varsul, que estuda a variação linguística na região Sul do Brasil. Trabalhamos com anotação automática de 20 textos, contendo 154.530 palavras. Este trabalho basicamente apresenta a ferramenta de anotação automática, o processo de análise morfossintática automática efetuada pelo anotador, nosso trabalho de revisão manual da etiquetagem automática e nossas sugestões de melhorias para tratar especificamente de aspectos da oralidade. Verificaremos o desempenho do Aelius na etiquetagem morfossintática automática de um *corpus* de língua falada, um desafio inédito para esta ferramenta. A partir dos erros do etiquetador, buscaremos depreender certos padrões de anotação para superar limitações apresentadas pelo programa, propondo algumas sugestões de implementações para que o Aelius etiquete de maneira ainda mais satisfatória *corpus* de língua falada. Trataremos especialmente dos casos de interjeições, aféreses, onomatopeias e marcadores conversacionais.

Palavras-chave: etiquetagem automática; etiquetagem morfossintática; linguística de corpus.

## ABSTRACT

This project is included in the area of Corpus Linguistics and morphological analysis of Brazilian Portuguese, which is about automatic morphological annotation *corpus* of spoken language. The work aims to contribute to the improvement of the automatic morphosyntactic tagger Aelius, originally developed by prof. Dr. Leonel Alencar, from Universidade Federal do Ceará, coordinator of the project CompLin - Natural Language and Computation. The tags used in Aelius are the same tags of Tycho-Brahe historical Portuguese *corpus* (electronic *corpus* already tagged, composed of texts in Portuguese written by authors born between 1380 and 1845). The texts that were used for tagging come from the bank of Varsul project, which studies the linguistic variation in southern Brazil. We worked with automatic tag of 20 texts, containing 154.530 words. This project presents the tool for automatic tagging, the automatic morphosyntactic analysis process executed by the tagger and our auto-tagging manual review work, besides some suggestions of improvement to deal specifically with aspects of orality. We will verify the Aelius' performance in automatic morphosyntactic tagging of a *corpus* of spoken language, an unprecedented challenge to this tool. From the errors of the tagger, we seek to infer certain patterns of annotation to overcome limitations presented by the program, proposing some suggestions of implementations so that Aelius can tag even more satisfactorily *corpus* of spoken language. We will deal especially with cases of interjections, apheresis, onomatopoeia and conversational markers.

Key-words: auto-tagging; morphosyntactic tagging; corpus linguistics.

## LISTA DE FIGURAS, QUADROS E TABELAS

Figura 1: Interface do etiquetador Aelius	19
Figura 2: Processo de anotação dos textos	24
Quadro 1: Exemplo de texto original (Varsul)	24
Quadro 2: Exemplo de texto “limpo”	25
Quadro 3: Exemplo de texto anotado automaticamente (Aelius)	25
Quadro 4: Exemplo de texto anotado automaticamente (Aelius) e revisado manualmente	26
Tabela 1: Constituição do <i>corpus</i>	22
Tabela 2: Acurácia do Aelius na nossa amostra do <i>corpus</i>	28
Tabela 3: Adjetivo etiquetado como nome	31
Tabela 4: Adjetivo etiquetado como nome próprio	32
Tabela 5: Adjetivo etiquetado como numeral	32
Tabela 6: Adjetivo etiquetado como verbo	33
Tabela 7: Advérbio etiquetado como adjetivo	33
Tabela 8: Advérbio etiquetado como nome	34
Tabela 9: Advérbio etiquetado como nome próprio	34
Tabela 10: Advérbio etiquetado como verbo	34
Tabela 11: Artigo etiquetado como clítico	34
Tabela 12: Artigo etiquetado como preposição	35
Tabela 13: Artigo etiquetado como verbo	35
Tabela 14: Conjunção etiquetada como preposição	35
Tabela 15: Interjeição etiquetada como adjetivo	35
Tabela 16: Interjeição etiquetada como nome	35
Tabela 17: Interjeição etiquetada como nome próprio	36
Tabela 18: Interjeição etiquetada como verbo	36
Tabela 19: Nome etiquetado como adjetivo	36
Tabela 20: Nome etiquetado como advérbio	38
Tabela 21: Nome etiquetado como clítico	38
Tabela 22: Nome etiquetado como nome próprio	39
Tabela 23: Nome próprio etiquetado como adjetivo	42

Tabela 24: Nome próprio etiquetado como nome comum	43
Tabela 25: Nome etiquetado como pronome demonstrativo	44
Tabela 26: Nome etiquetado como pronome interrogativo	44
Tabela 27: Nome etiquetado como verbo	44
Tabela 28: Nome etiquetado como verbo + clítico	46
Tabela 29: Nome plural terminado em -S etiquetado como nome singular	47
Tabela 30: Nome singular terminado em -S etiquetado como nome plural	47
Tabela 31: Numeral etiquetado como adjetivo	47
Tabela 32: Numeral etiquetado como clítico	47
Tabela 33: Numeral etiquetado como pronome definido	47
Tabela 34: Pronome demonstrativo etiquetado como nome próprio	48
Tabela 35: Pronome de tratamento etiquetado como nome próprio	48
Tabela 36: Pronome de tratamento etiquetado como pronome possessivo	48
Tabela 37: Pronome pessoal etiquetado como nome plural	48
Tabela 38: Pronome pessoal etiquetado como nome próprio	49
Tabela 39: Pronome pessoal etiquetado como preposição + demonstrativo	49
Tabela 40: Verbo etiquetado como adjetivo	50
Tabela 41: Verbo etiquetado como advérbio	50
Tabela 42: Verbo etiquetado como nome	51
Tabela 43: Verbo etiquetado como nome próprio	51
Tabela 44: Verbo conjugado etiquetado como verbo infinitivo	51
Tabela 45: Verbo conjugado etiquetado como <i>SER</i> verbo infinitivo	52
Tabela 46: Verbo terminado em <i>-rei</i> etiquetado como verbo futuro	52
Tabela 47: Verbo terminado em <i>-S</i> etiquetado como nome plural	53
Tabela 48: Verbo etiquetado como preposição	53
Tabela 49: Verbo presente etiquetado como verbo gerúndio	53
Tabela 50: Verbo presente etiquetado como verbo imperativo	54
Tabela 51: Verbo presente etiquetado como verbo infinitivo	54
Tabela 52: Verbo indicativo etiquetado como verbo subjuntivo	54
Tabela 53: Verbo gerúndio etiquetado como verbo particípio	54
Tabela 54: Verbo presente etiquetado como particípio	54
Tabela 55: Verbo passado etiquetado como verbo presente	55
Tabela 56: Erros de anotação da fala	56





## SUMÁRIO

<b>1 INTRODUÇÃO</b>	11
<b>2 FUNDAMENTAÇÃO TEÓRICA</b>	13
2.1 O que é Linguística de Corpus	13
2.2 O que é um <i>corpus</i>	14
2.2.1 Como analisar um <i>corpus</i>	15
2.2.2 Etiquetagem automática morfossintática	16
2.2.3 Etiquetagem automática morfossintática de língua falada	17
<b>3 A FERRAMENTA E O CORPUS</b>	19
3.1 Aelius	19
3.2 Varsul	20
<b>4 ANÁLISE</b>	23
4.1 Processamento dos textos	23
4.2 Exemplos de textos	24
4.3 Correção manual da etiquetagem automática	26
4.3.1 Principais problemas encontrados na correção manual	29
4.3.1.1 Erros de “confusão” de etiquetas	30
4.3.1.2 Erros de anotação da fala	55
<b>5 SUGESTÕES PARA IMPLEMENTAÇÕES FUTURAS NO AELIUS</b>	60
5.1 Sugestões e novas etiquetas	60
5.1.1 Inserção de palavras na etiqueta interjeição	61
5.1.2 Inserção de nova etiqueta	62
5.1.2.1 Onomatopeias	62
5.1.2.2 Aférese	63
5.1.2.3 Marcadores conversacionais	63
<b>6 CONSIDERAÇÕES FINAIS</b>	66

<b>7 REFERÊNCIAS</b>	68
<b>8 ANEXOS</b>	70
7.1 Texto cru	70
7.2 Texto anotado automaticamente pelo Aelius	75
7.3 Etiquetas	87

## 1 INTRODUÇÃO

Com o avanço da tecnologia, temos melhorias em vários ramos da ciência, e com a Linguística não poderia ser diferente. Portanto, esta pesquisa pretende contribuir com a melhoria de um programa de anotação morfossintática automática, o etiquetador automático morfossintático Aelius. Temos o objetivo de analisar este anotador automático, que é gratuito e robusto, a fim de que ele possa dar conta de etiquetar *corpora* de língua falada, além de *corpora* de língua escrita, que ele já etiqueta de maneira satisfatória. Nossa intenção última é aprimorar o Aelius para que ele possa ser utilizado para anotação automática de *corpora* falado, em especial, o *corpus* coletado pelo Varsul, tendo em vista que os textos anotados automaticamente podem auxiliar várias pesquisas linguísticas, tanto de cunho teórico como aplicado, de acordo com Lemnitzer e Zinsmeister (2006), *apud* Alencar (2009) “os *corpora* anotados automaticamente têm se tornado cada vez mais indispensáveis à pesquisa não só em linguística descritiva, mas também teórica”.

Este trabalho de conclusão é o ápice de meu trabalho como bolsista de iniciação científica, cujos resultados parciais já foram apresentados na III Jornada do Varsul e no XXV Salão de Iniciação Científica da Universidade Federal do Rio Grande do Sul. Nossa pesquisa insere-se duplamente nos projetos Variação Linguística na Região Sul do Brasil - VARSUL<sup>1</sup> e Computação e Linguagem Natural - Complín<sup>2</sup>, do qual meu orientador é membro integrante. O trabalho está estruturado e organizado da maneira que segue: a seção 2 discute a fundamentação teórica do trabalho, percorrendo um pouco sobre o que é a Linguística de Corpus, o que é um *corpus* e como ele pode ser analisado, com enfoque à etiquetagem morfossintática automática de textos, que é o assunto principal deste trabalho; a seção 3 apresenta a ferramenta computacional e o *corpus* utilizados na pesquisa, ou seja, o etiquetador Aelius e o *corpus* de transcrições de entrevistas do Varsul; a seção 4 traz a análise da correção manual da anotação automática feita pelo Aelius, onde será descrita como foi feita a revisão manual da anotação automática e serão apresentadas as principais dificuldades encontradas, além da esquematização dos equívocos cometidos pelo Aelius; a seção 5 discute sugestões de implementações para melhoria do etiquetador Aelius em relação à anotação automática de *corpus* de língua falada, tendo como base a revisão manual feita sobre a etiquetagem

---

<sup>1</sup> <http://www.varsul.org.br/>

<sup>2</sup> <http://complin.blogspot.com.br/>

automática de nossa amostra do *corpus* do Varsul; finalmente, a seção 6 traz algumas considerações finais sobre este trabalho. No Apêndice, o leitor poderá consultar um texto anotado automaticamente pelo Aelius.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo abordaremos questões iniciais, do que trata Linguística de Corpus, como se caracteriza um *corpus* linguístico, e quais as serventias de um *corpus* – e de ferramentas computacionais que trabalham com corpus – para a pesquisa linguística atualmente.

### 2.1 O que é Linguística de Corpus

A Linguística de Corpus trata de coletar, compilar e explorar conjuntos de textos para pesquisa linguística de uma determinada língua, ou de comparação entre duas ou mais línguas, ou ainda, de uma variedade linguística – e é importante que essa amostra de língua esteja em um formato que seja processável pelo computador. Essa vertente do estudo da linguagem recebeu mais força quando estudiosos buscaram alternativas para se apoiar no uso real da língua, tendo em vista que os *corpora* permitem que se pesquise e se baseie em linguagem natural e autêntica, possibilitando, assim, a comprovação de teorias por meio de análise dos dados efetivamente produzidos por falantes da língua. Quando se trata de Linguística de Corpus, tanto a teoria quanto a prática são extremamente necessárias, é praticamente impossível pensar Linguística de Corpus sem praticá-la, sem fazê-la. De acordo com Berber Sardinha (2000b: 2),

a Linguística de Corpus se ocupa da coleta e exploração de *corpora*, ou conjuntos de dados linguísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem através de evidências empíricas, extraídas por meio de computador.

Atualmente, os *corpora* são armazenados em computadores que têm bastante memória HD (para armazenamento de textos) e memória RAM (para velocidade no processo de manuseio de textos), o que possibilita armazenar uma grande quantidade de

informações, diferentemente dos primeiros *corpora* que surgiram, escritos à mão por um estudioso ou por um grupo de pessoas – como compilações de textos bíblicos, por exemplo. O aumento dos estudos que utilizam ou se relacionam a *corpora* linguísticos aumentou – especialmente desde a década de 1990 para cá – devido ao aperfeiçoamento da tecnologia dos computadores, que permite maior visibilidade aos *corpora* linguísticos para os pesquisadores, além de garantir mais eficiência em tarefas complexas. Importante destacar que há também um grande desenvolvimento na área da Linguística de Corpus, financiado por empresas, com intenções comerciais, como, por exemplo, reconhecedores de fala e informatização de grandes bases de dados.

Tendo isso em vista, pode-se notar claramente que, conforme a tecnologia avança, a Linguística de Corpus é aperfeiçoada, também devido a algumas ferramentas computacionais que são criadas para um melhor aproveitamento dos *corpora*, como etiquetadores automáticos, concordanciadores, tradutores automáticos, etc. É evidente que os linguistas de *corpus* necessitam de programas que permitam explorar os *corpora* e extrair as informações das quais precisam em suas pesquisas; por esse motivo, novos programas surgem e são constantemente aperfeiçoados. A área da Linguística de Corpus não chega a ser uma área independente, mas serve como uma própria ferramenta aos estudos linguísticos que almejam se basear em dados empíricos.

## 2.2 O que é um corpus

Um *corpus* é um conjunto de textos ou dados linguísticos autênticos, coletados e selecionados com uma intenção de pesquisa, de maneira que se parte da motivação da pesquisa para o corpus e não o contrário. Um *corpus* deve representar determinada variedade linguística, que servirá como base para análise e pesquisa. Segundo Sanchez e Cantos (1996), *apud* Sardinha (2004), um corpus é:

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador,

com a finalidade de propiciar resultados vários e úteis para a descrição e análise.

Para que um *corpus* seja estudado, é preciso que o pesquisador delimite seu enfoque de estudo, para então definir qual é o corpus mais adequado para seu objetivo de pesquisa. O corpus, independentemente de sua dimensão, deve ser representativo de uma língua ou de uma variedade dela. É evidente que, quanto maior for o *corpus*, mais representativo ele será, e isso se deve a dois fatores: primeiramente porque “a linguagem é um sistema probabilístico, no qual certos traços são mais frequentes que outros”, e depois porque “o *corpus* é uma amostra de uma população cuja dimensão não se conhece (a linguagem como um todo)” (Berber Sardinha, 2000b: 5).

Um *corpus* eletrônico é feito para ser usado por pessoas (o que diferencia a Linguística de Corpus do Processamento de Dados), ou seja, um linguista de *corpus*, antes de qualquer coisa é um linguista, e carrega consigo questões específicas que norteiam seu interesse na exploração do *corpus*. Muitas das informações que o linguista quer buscar no *corpus* não estão explícitas; por exemplo, um *corpus* não reconhece o que é “sujeito” ou o que é “verbo”. Esses subsídios precisam ser fornecidos, para que as perguntas que um linguista possa fazer possam ser respondidas (como por exemplo, quais os adjetivos usados no *corpus*, etc.). O fornecimento dessa informação é o que chamamos de **anotação**, que é o principal assunto deste trabalho.

### 2.2.1 Como analisar um corpus

Além da necessidade de escolher bem o *corpus* para sua pesquisa, o pesquisador precisa escolher que ferramenta lhe pode ser útil para analisar o *corpus*, dependendo de seu objetivo. Vale destacar, novamente, a importância do computador para tarefas que poderiam ser muito exaustivas e levar um longo tempo se fossem feitas com nossas limitações humanas, como por exemplo, contar palavras. Existem muitas ferramentas que têm capacidade para fazer muitas operações; o que falta, talvez, é conhecimento por parte dos linguistas para utilizar essas ferramentas. Muitos nem as conhecem, outros já ouviram falar, mas não se interessaram por buscar saber usá-las efetivamente. Outro

ponto é que podemos usar uma ferramenta que foi criada para devida finalidade para um outro desígnio, sabendo aproveitar a tecnologia que está a nosso favor.

Há ferramentas para diversas finalidades: ferramentas que medem a dificuldade de compreensão de um texto (índice Flesch), que contam o número de palavras, sentenças, parágrafos, palavras por sentenças, sentenças por parágrafos, sílabas por palavras, incidência de verbos, substantivos, adjetivos, advérbios, pronomes, palavras de conteúdo, palavras funcionais, etc.. Há ferramentas que criam listas de palavras, conforme sua frequência ou em ordem alfabética, ferramentas que criam lista de palavras-chave, concordanciadores, etc..

Além disso, há a etiquetagem, que pode ser morfossintática (*part of speech*), que trata de classificar as unidades lexicais do texto através de *tags* (etiquetas) morfossintáticas pré-estabelecidas em uma lista de etiquetas; etiquetagem sintática (*parsing*), que identifica no texto as estruturas sintáticas presentes, os sintagmas; etiquetagem semântica, que trata especialmente dos itens lexicais que podem apresentar mais de um sentido em um determinado contexto; e, por fim, a etiquetagem discursiva, que traz aspectos relacionados à Linguística do Texto e à Análise do Discurso, marcando as palavras como tópicos ou referentes anafóricos, por exemplo<sup>1</sup>.

Neste trabalho, trataremos especificamente da etiquetagem morfossintática automática, que marca após cada palavra do *corpus* qual é sua classe gramatical, baseando-se num conjunto de etiquetas morfossintáticas pré-estabelecida.

### 2.2.2 Etiquetagem automática morfossintática

A etiquetagem de *corpus* é uma parte muito importante nas pesquisas com *corpora*, pois é a etiquetagem que vai permitir que o *corpus* cumpra seu papel como instrumento para investigação. A etiquetagem morfossintática permite que uma ferramenta, através do computador, etiquete uma grande quantidade de texto de maneira eficiente. De acordo com Alencar (2012), a etiquetagem “é uma tarefa aparentemente

---

<sup>1</sup> Sobre as diversas aplicações da Linguística de Corpus, remetemos o leitor a Garside, Leech & McEnery (1997), Othero & Gasperin (2001), Berber Sardinha, 2000a, 2000b, 2004.



simples para o processamento da linguagem, no entanto, o desempenho de outras ferramentas depende diretamente desse processo”. Além disso, a importância do anotador automático também diz respeito à questão de se entender a linguagem natural, “a etiquetagem morfossintática é uma tarefa intermediária que tem como objetivo principal analisar e entender a língua natural” (ALENCAR, 2012).

Para que um texto seja etiquetado, é necessário primeiramente fazer sua tokenização, ou seja, dividir o texto em *tokens*, que podem ser sentenças ou palavras. Além disso, é indispensável que o texto seja “limpado”, retirando todas as marcas que podem atrapalhar a etiquetagem, como por exemplo, apóstrofes e travessões. Essa limpeza do texto também pode ser feita de maneira automática, utilizando-se de um *script*, que é um código programado para eliminar essas marcações automaticamente. Não faz parte das expectativas do anotador receber um texto com tantas marcações; por isso é de extrema importância que o texto esteja limpo, para maior eficiência por parte do etiquetador.

Então, sobre o texto cru, original, é feita a tokenização e a limpeza; assim, o texto estará pronto para passar pela fase de etiquetagem, em que cada *token* recebe uma etiqueta conforme sua classe gramatical - e de acordo com as etiquetas morfossintáticas disponíveis pelo etiquetador.

### **2.2.3 Etiquetagem automática morfossintática de língua falada**

Ao longo dos últimos anos, a Linguística de Corpus no Brasil aumentou seu escopo de pesquisa, abrangendo a língua falada além da língua escrita. Sobre a importância de dar uma atenção especial e diferenciada à língua falada, podemos citar dois principais motivos: primeiro, todas as línguas possuem oralidade, mas nem todas possuem escrita; sendo assim, podemos, por exemplo, coletar e organizar *corpora* de línguas indígenas que não possuem ortografia, para pesquisá-las de maneira mais concreta. Em segundo lugar, a escrita não figura como uma representação da fala, pois não mostra alguns dos fenômenos da oralidade, como por exemplo, contrações e truncamentos.

Usualmente, as transcrições de *corpora* falados possuem essas características da fala, como contrações, fragmentos de sentenças, interjeições e truncamentos. Por isso, é de extrema importância que o etiquetador dê conta de etiquetar corretamente todas essas informações que fazem parte da fala e não podem ser ignoradas. Além disso, é importante que essas ocorrências recebam a etiqueta correta porque, como as etiquetas se relacionam umas com as outras, ao receber uma equivocada, uma palavra pode ocasionar na anotação equivocada de outras palavras próximas a ela. É o que veremos nos próximos capítulos com a etiquetagem automática de trechos do *corpus* de língua falada do Varsul pelo etiquetador Aelius.

### 3 A FERRAMENTA E O CORPUS

Neste capítulo, apresentaremos a ferramenta de anotação automática utilizada, o etiquetador Aelius, e o *corpus* de língua falada usado para anotação, trechos extraídos do banco de dados do projeto Varsul.

#### 3.1 O Aelius

O projeto do etiquetador Aelius surgiu da necessidade de tornar acessível a estudantes e pesquisadores de Linguística a análise automática de textos, além de ser uma maneira de iniciar estudos na área da Linguística Computacional, por parte de alunos de Engenharia, Computação e Letras. O etiquetador Aelius foi desenvolvido pelo professor Leonel Alencar, da Universidade Federal do Ceará, que coordena o grupo CompLin – Computação e Linguagem Natural, registrado no CNPq.

O Aelius é um *software* que pode ser baixado no site <http://aelius.sourceforge.net/manual.html>, e apresenta uma interface relativamente amigável, como podemos notar abaixo:

```

1  >>> from Aelius import AnotaCorpus
2  >>> from Aelius import ProcessaNomesProprios
3  >>> sentes=ProcessaNomesProprios.SENTENCAS
4  >>> tokens=[sent.split() for sent in sentes]
5  >>> codificadas=AnotaCorpus.codifica_sentencas(tokens)
6  >>> for s in rubt.batch_tag(codificadas[:3]):
7     for w,t in s:
8         print "%s/%s " % (w,t),
9     print
10
11
12 -( Luzia/NPR pediu/VB-D a/P Deus/NPR e/CONJ a/D-F Ávila/NPR para/P que/WPRO lhe/CL ajudassem/VB-SD a/D-F sair/VB de/P Sobral
13 Deus/NPR ajudou/VB-D Luzia/NPR ./
14 .../ Sobral/NPR era/SR-D uma/D-UM-F cidade/N intelectual/ADJ-G ./
15 .../ Cidade/N intelectual/ADJ-G ,/ Sobral/NPR tinha/TR-D muitos/Q-P poetas/N-P ./
16 Municipio/NPR intelectual/ADJ-G ,/ Sobral/NPR tinha/TR-D muitos/Q-P poetas/N-P ./
17 Fortaleza/NPR era/SR-D uma/D-UM-F cidade/N provinciana/ADJ-F ./
18 Ávila/NPR ajudou/VB-D Luzia/NPR ./
19 .../ Cansada/VB-AN-F ,/ Luzia/NPR logo/ADV dormiu/VB-D ./
20 Ávida/N por/P sossego/N ,/ Luzia/NPR deixou/VB-D a/P cidade/N ./
21 Ótimo/N !/
22 Bom/ADJ !/
23 .../ -( Bom/ADJ ./

```

Figura 1: Interface do etiquetador automático Aelius

O Aelius é uma ferramenta para analisar textos automaticamente, contando com as etiquetas utilizadas no *corpus* de português histórico Tycho-Brahe “(*corpus* eletrônico já anotado, composto de textos em português escritos por autores nascidos entre 1380 e

1845)<sup>1</sup>. Atualmente, 57 textos (2.547.503 palavras) estão disponíveis para pesquisa livre, com um sistema de anotação linguística em duas etapas: **anotação morfológica** (aplicada em 33 textos) e **anotação sintática** (aplicada em 16 textos)” (ALENCAR, 2012). Esse *corpus* é desenvolvido junto ao projeto temático Padrões Rítmicos, Fixação de Parâmetros & Mudança Linguística da Universidade Estadual de Campinas<sup>2</sup>. Segundo Alencar (2010), “o nível alcançado de acurácia da etiquetagem supera o de ferramentas análogas livremente disponíveis, voltadas sobretudo para o português contemporâneo”, ou seja, dentre as ferramentas que temos disponíveis, o Aelius é uma ótima alternativa para a etiquetagem de *corpora*, pois tem um alto índice de acerto.

O nome desse etiquetador é uma homenagem ao gramático latino Aelius Donatus, que foi o gramático mais influente de seu tempo. Foi ele quem analisou e definiu uma série de palavras como nomes, verbos, preposições, etc. De acordo com Dezotti (2010: 2), Aelius Donatus, com sua obra *Ars Donati Grammatici Urbis Romae*, “suplantou todas as outras e alterou o modo de produção de manuais de gramática no fim da Antiguidade”. A primeira parte da obra de Donatus trata das oito partes discurso, e a segunda parte trata de fonética, métrica e estilística.

### 3.2 Varsul

O Variação Linguística na Região Sul do Brasil – Varsul é um projeto que estuda a variação linguística na região Sul do Brasil e conta com a parceria de quatro universidades brasileiras, a saber: Universidade Federal do Rio Grande do Sul, Pontifícia Universidade Católica do Rio Grande do Sul, Universidade Federal de Santa Catarina e Universidade Federal do Paraná.

Esse projeto foi criado no ano de 1982, seguindo a proposta da professora Leda Bisol, e tem como objetivos principais proporcionar subsídios para descrever a língua

---

<sup>1</sup> Todas as etiquetas utilizadas pelo Aelius (que foram emprestadas do projeto Tycho Brahe) podem ser consultadas no apêndice deste trabalho.

<sup>2</sup> <http://www.tycho.iel.unicamp.br/~tycho/prfpml/fase2/>

falada no Brasil; criar condições para teste e desenvolvimento de teorias linguísticas; e oferecer subsídios para programas educacionais, promovendo, assim, o conhecimento e o respeito às variedades linguísticas<sup>3</sup>.

Atualmente o projeto do Varsul se divide em três: Banco de Dados do VARSUL, Amostra Digital VARSUL e Banco de Dados Diacrônico<sup>4</sup>. Segundo Mello e Raso (2012), nas entrevistas há três tipologias de interação: a monológica, que ocorre quando um falante sozinho fala sem interação; a dialógica, onde dois falantes conversam entre si; e, a tipologia conversacional, que ocorre quando dois ou mais falantes interagem. Todas as entrevistas que analisamos foram de interação dialógica, registrando um diálogo entre um aluno da UFRGS, bolsista do Varsul, e a pessoa entrevistada.

Em nossa pesquisa, analisamos, automática e manualmente, como veremos, vinte trechos de entrevistas transcritas do Banco de Dados do Varsul, todas ocorridas em Porto Alegre, entre os anos de 1990 e 1999. Esses vinte trechos totalizam 410 páginas e 154.530 palavras. Na tabela abaixo, temos esquematizado nosso *corpus*, apresentando o nome de referência de cada trecho do Varsul, além de seu número de páginas e de palavras.

---

<sup>3</sup> Ver, por exemplo, os trabalhos Schwindt (1995), Tavares (1999), Coelho (2000) e Mittmann (2006), entre outros.

<sup>4</sup> As entrevistas do Varsul estão disponíveis nas universidades participantes do projeto.

<b>REFÊRENCIA VARSUL</b>	<b>Nº de PÁGINAS</b>	<b>Nº de PALAVRAS</b>
<b>POA 01</b>	<b>7</b>	<b>3.338</b>
<b>POA 01.1</b>	<b>23</b>	<b>9.572</b>
<b>POA 02</b>	<b>8</b>	<b>3.285</b>
<b>POA 02.1</b>	<b>31</b>	<b>11.572</b>
<b>POA 03</b>	<b>8</b>	<b>2.965</b>
<b>POA 03.1</b>	<b>27</b>	<b>9.919</b>
<b>POA 04</b>	<b>17</b>	<b>6.877</b>
<b>POA 05</b>	<b>7</b>	<b>2.923</b>
<b>POA 12</b>	<b>8</b>	<b>2.797</b>
<b>POA 15</b>	<b>8</b>	<b>2.691</b>
<b>POA 21</b>	<b>9</b>	<b>3.338</b>
<b>POA 25</b>	<b>29</b>	<b>10.890</b>
<b>POA 26</b>	<b>28</b>	<b>10.152</b>
<b>POA 27</b>	<b>30</b>	<b>10.426</b>
<b>POA 28</b>	<b>29</b>	<b>10.510</b>
<b>POA 29</b>	<b>30</b>	<b>11.908</b>
<b>POA 31</b>	<b>29</b>	<b>11.467</b>
<b>POA 32</b>	<b>30</b>	<b>11.880</b>
<b>POA 35</b>	<b>18</b>	<b>7.397</b>
<b>POA 41</b>	<b>34</b>	<b>10.623</b>
<b>TOTAL</b>	<b>410</b>	<b>154.530</b>

**Tabela 1:** Constituição do *corpus*

A partir desse corpus e com a ferramenta apresentada, partiremos para as análises da etiquetagem automática morfossintática, no próximo capítulo, esquematizando os principais problemas encontrados, mostrando os contextos em que ocorreram.

## 4 ANÁLISE

Neste capítulo, falaremos sobre o processo prático da pesquisa – a revisão manual do *corpus* anotado automaticamente - e como se deu a análise do *corpus* anotado.

### 4.1 Processamento dos textos

Primeiramente, foram selecionados alguns trechos de transcrições de entrevistas do banco de dados de Porto Alegre do Varsul, aleatoriamente. Selecionados os textos, que estavam em formato .doc, foi necessário passá-los para o formato .txt, tendo em vista que é o formato que o etiquetador Aelius espera receber seus arquivos para poder rodar os textos e fazer a etiquetagem.

Esses trechos foram enviados por e-mail para o professor Leonel e seu grupo de pesquisa da UFC, que usaram o *script* adequado para limpar automaticamente os textos, retirando todas as marcas que poderiam atrapalhar a anotação, como travessões, asteriscos e parênteses, por exemplo. Após os textos estarem “limpos”, foram rodados no etiquetador automático Aelius, e depois, enviados de volta para nós, via e-mail. Por último, os textos foram impressos e revisados manualmente<sup>1</sup>.

---

<sup>1</sup> É possível verificar, no apêndice, uma amostra do texto cru e do texto anotado pelo Aelius.

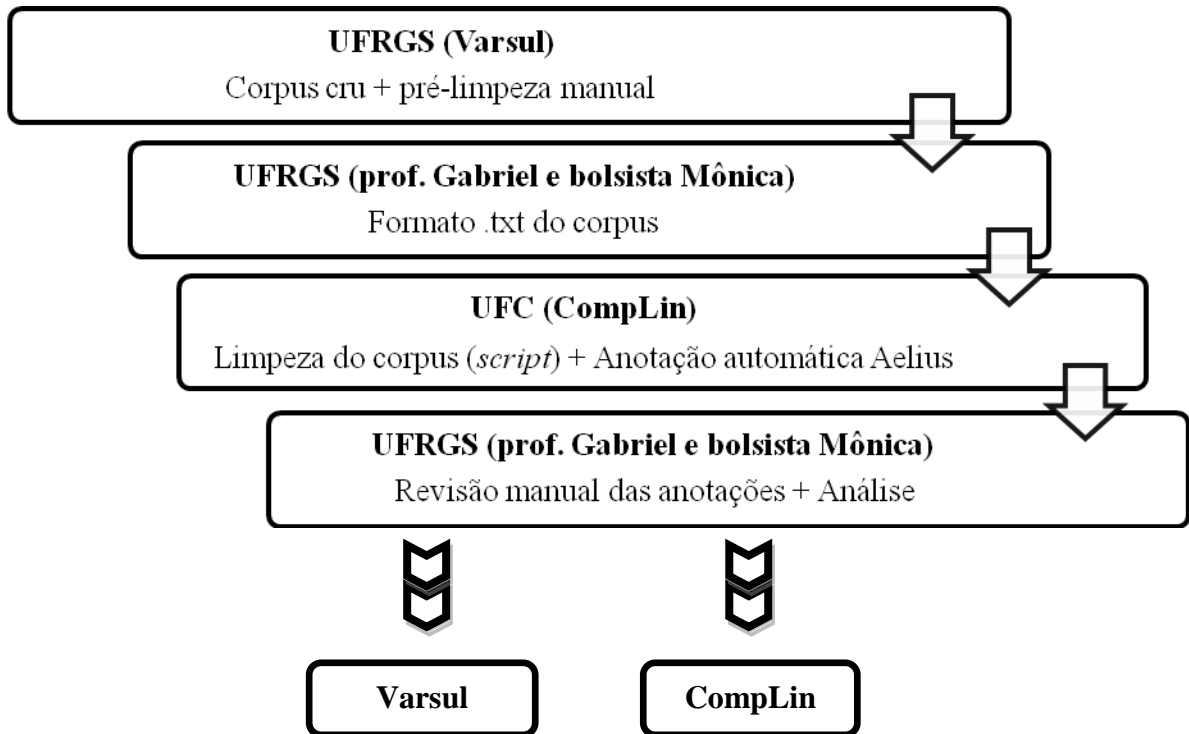


Figura 2: processo de anotação dos textos

#### 4.2 Exemplos de textos:

Neste primeiro quadro, trazemos um trecho como exemplo de texto cru, em que notamos que há caracteres que podem atrapalhar o etiquetador, como asteriscos, sustenidos, e colchetes.

---

... entregar pra amanhã, vai se entregar, se Deus quiser, salvo algum problema, né? extra, ma0s está tudo bem, vai ser entregue, se Deus quiser. \*Então, [isso]-[isso]- sabe que isso é muito importante, sabe o que que é? \*Então a pessoa fa0z uma previsão: "\*Eu vou entregar o serviço-" \*Pra ti no caso: "\*Olha, eu vou entregar teu carro sexta#feira." \*Se eu não entregar sexta#feira vai ficar meio ridículo, meio xarope, sabe? \*Você vai entender, ma0s no fundo, no fundo você: "\*Pô ma0s Eduardo0 não cumpriu com a palavra." \*No seu íntimo, né?...

---

Quadro 1: Exemplo de texto original (Varsul)

O segundo quadro é uma amostra de como o texto fica após ser “limpo” pelo *script*, sem as marcas que podem interferir na etiquetagem.



---

... entregar pra amanhã, vai se entregar, se Deus quiser, salvo algum problema, né? extra, mas está tudo bem, vai ser entregue, se Deus quiser. Então, isso isso sabe que isso é muito importante, sabe o que que é? Então a pessoa faz uma previsão: "Eu vou entregar o serviço" Pra ti no caso: "Olha, eu vou entregar teu carro sexta-feira." Se eu não entregar sexta-feira vai ficar meio ridículo, meio xarope, sabe? Você vai entender, mas no fundo, no fundo você: "Pô mas Eduardo não cumpriu com a palavra." No seu íntimo, né?...

---

**Quadro 2:** Exemplo de texto "limpo"

No terceiro quadro, temos um exemplo de como o texto fica após ser etiquetado automaticamente pelo Aelius. Ao ser anotada, cada palavra recebe um símbolo / (barra) e após esse símbolo uma etiqueta que define sua classe, como podemos observar abaixo.

---

... entregar/VB pra/P amanhã/ADV ,/, vai/VB-P se/SE entregar/VB-SR ,/, se/CONJS Deus/NPR quiser/VB-SR ,/, salvo/P algum/Q problema/N ,/, né/NPR ?/. Extra/ADV ,/, mas/CONJ está/ET-P tudo/Q bem/ADV ,/, vai/VB-P ser/SR entregue/VB-AN ,/, se/CONJS Deus/NPR quiser/VB-SR ./.. Então/ADV ,/, isso/DEM isso/DEM sabe/VB-P que/C isso/DEM é/SR-P muito/Q importante/ADJ-G ,/, sabe/VB-P o/D que/WPRO que/WPRO é/SR-P ?/. Então/ADV a/D-F pessoa/N faz/VB-P uma/D-UM-F previsão/N :/. "/QT Eu/PRO vou/VB-P entregar/VB o/D serviço/N "/QT Pra/P ti/PRO no/P+D caso/N :/. "/QT Olha/VB-P ,/, eu/PRO vou/VB-P entregar/VB teu/PRO\$ carro/N sexta-feira/NPR ./.. "/QT Se/CONJS eu/PRO não/NEG entregar/VB sexta-feira/N vai/VB-P ficar/VB meio/N ridículo/ADJ ,/, meio/ADJ xarope/N ,/, sabe/VB-P ?/. Você/PRO vai/VB-P entender/VB ,/, mas/CONJ no/P+D fundo/N ,/, no/P+D fundo/N você/PRO :/. "/QT Pô/NPR mas/CONJ Eduardo/NPR não/NEG cumpriu/VB-D com/P a/D-F palavra/N ./.. "/QT No/P+D seu/PRO\$ íntimo/N ,/, né/NPR ?/. ...

---

**Quadro 3:** Exemplo de texto anotado automaticamente (Aelius)<sup>2</sup>

<sup>2</sup> Uma lista contendo todas as etiquetas poderá ser consultada no apêndice deste trabalho.

### 4.3 Correção manual da etiquetagem automática

A correção manual da etiquetagem automática é um trabalho que exige muita atenção e revisão, pois são muitas as etiquetas que podem aparecer, e, uma mesma palavra pode aparecer com etiquetas diferentes, dependendo dos contextos em que estiver inserida. Por isso, deve-se prestar atenção ao contexto da palavra e não apenas à etiqueta da palavra isolada ou pensada em um contexto abstrato.

Neste quadro, podemos ter uma ideia de como se dá a revisão manual da etiquetagem automática, pois aqui trazemos um exemplo de texto que foi devidamente “limpo”, etiquetado automaticamente e revisado manualmente. As etiquetas marcadas equivocadamente estão destacadas em negrito.

---

... entregar/VB pra/P amanhã/ADV ,/, vai/VB-P se/SE **entregar/VB-SR** ,/, se/CONJS Deus/NPR **quiser/VB-SR** ,/, **salvo/P** algum/Q problema/N ,/, **né/NPR?/**. **Extra/ADV** ,/, mas/CONJ está/ET-P tudo/Q bem/ADV ,/, vai/VB-P ser/SR entregue/VB-AN ,/, se/CONJS Deus/NPR quiser/VB-SR ./ Então/ADV ,/, isso/DEM isso/DEM sabe/VB-P que/C isso/DEM é/SR-P muito/Q importante/ADJ-G ,/, sabe/VB-P o/D que/WPRO que/WPRO é/SR-P ?/. Então/ADV a/D-F pessoa/N faz/VB-P uma/D-UM-F previsão/N :/. "/QT Eu/PRO vou/VB-P entregar/VB o/D serviço/N "/QT Pra/P ti/PRO no/P+D caso/N :/. "/QT Olha/VB-P ,/, eu/PRO vou/VB-P entregar/VB teu/PRO\$ carro/N **sextafeira/NPR** ./ "/QT Se/CONJS eu/PRO não/NEG entregar/VB sextafeira/N vai/VB-P ficar/VB **meio/N** ridículo/ADJ ,/, meio/ADJ **xarope/N** ,/, sabe/VB-P ?/. Você/PRO vai/VB-P entender/VB ,/, mas/CONJ no/P+D fundo/N ,/, no/P+D fundo/N você/PRO :/. "/QT **Pô/NPR** mas/CONJ Eduardo/NPR não/NEG cumpriu/VB-D com/P a/D-F palavra/N ./ "/QT No/P+D seu/PRO\$ íntimo/N ,/, **né/NPR?/** ...

---

**Quadro 4:** Exemplo de texto anotado automaticamente (Aelius) e revisado manualmente

Para fazer a correção manual desse *corpus*, analisamos os 20 textos destacando as etiquetas que apareceram erradas, e após essa primeira análise, revisamos os textos para garantir que nenhuma etiqueta passasse despercebida.

A tabela 2, a seguir, ilustra a composição do *corpus* e a acurácia das análises do anotador. Na primeira coluna, referência Varsul, temos o nome do arquivo que nos foi fornecido do banco de dados do Varsul, que são os nomes das transcrições das entrevistas. Nas linhas em que temos o mesmo nome de referência, foram analisados dois trechos da mesma entrevista. Na segunda coluna, temos a informação do número de palavras totais de cada trecho e na coluna seguinte, o número de palavras que foram etiquetadas erradamente. Na terceira coluna, calculamos a acurácia do anotador em cada um dos trechos individualmente. Na última linha temos a totalização dos números apresentados na tabela, que são os números totais do *corpus* que analisamos<sup>3</sup>, contendo 154.530 palavras, das quais 7.004 foram etiquetadas equivocadamente, de maneira que a acurácia do Aelius é de 95,4% em nosso *corpus*.

---

<sup>3</sup> Para fazer a correção manual desses trechos, imprimimos todos eles, e analisamos que etiqueta cada palavra recebeu, conferindo na lista de etiquetas. No começo era necessário nos remetermos à lista em cada palavra, mas após os primeiros trechos, já conseguimos decorar a maioria das etiquetas e a consulta à lista diminuiu, ocorrendo apenas em casos de etiquetas menos frequentes. Após essa primeira análise, revisamos mais uma vez cada texto, conferindo se não havíamos esquecido ou ignorado nenhuma ocorrência.

<b>REFÊRENCIA VARSUL</b>	<b>Nº de PALAVRAS</b>	<b>Nº de ERROS</b>	<b>ACURÁCIA</b>
<b>POA 01</b>	<b>3.338</b>	<b>203</b>	<b>93,9%</b>
<b>POA 01.1</b>	<b>9.572</b>	<b>508</b>	<b>94,6%</b>
<b>POA 02</b>	<b>3.285</b>	<b>135</b>	<b>95,8%</b>
<b>POA 02.1</b>	<b>11.572</b>	<b>398</b>	<b>96,5%</b>
<b>POA 03</b>	<b>2.965</b>	<b>96</b>	<b>96,7%</b>
<b>POA 03.1</b>	<b>9.919</b>	<b>351</b>	<b>96,4%</b>
<b>POA 04</b>	<b>6.877</b>	<b>252</b>	<b>96,3%</b>
<b>POA 05</b>	<b>2.923</b>	<b>108</b>	<b>96,3%</b>
<b>POA 12</b>	<b>2.797</b>	<b>121</b>	<b>95,6%</b>
<b>POA 15</b>	<b>2.691</b>	<b>189</b>	<b>92,9%</b>
<b>POA 21</b>	<b>3.338</b>	<b>180</b>	<b>94,6%</b>
<b>POA 25</b>	<b>10.890</b>	<b>466</b>	<b>95,7%</b>
<b>POA 26</b>	<b>10.152</b>	<b>419</b>	<b>95,8%</b>
<b>POA 27</b>	<b>10.426</b>	<b>337</b>	<b>96,7%</b>
<b>POA 28</b>	<b>10.510</b>	<b>456</b>	<b>95,6%</b>
<b>POA 29</b>	<b>11.908</b>	<b>532</b>	<b>95,5%</b>
<b>POA 31</b>	<b>11.467</b>	<b>824</b>	<b>92,8%</b>
<b>POA 32</b>	<b>11.880</b>	<b>432</b>	<b>96,3%</b>
<b>POA 35</b>	<b>7.397</b>	<b>531</b>	<b>92,8%</b>
<b>POA 41</b>	<b>10.623</b>	<b>466</b>	<b>95,6%</b>
<b>TOTAL</b>	<b>154.530</b>	<b>7.004</b>	<b>95,4%</b>

**Tabela 2:** Acurácia do Aelius na nossa amostra do *corpus*

Percebemos que mesmo antes da nossa análise e de nossa proposta de novas etiquetas, o Aelius tem uma acurácia considerada alta (95,4%). Entretanto, sua eficiência ainda é maior na análise de textos escritos: 96,3% de acurácia (cf. ALENCAR, 2013). Para termos uma noção, podemos compará-lo ao etiquetador LX-Tagger, que tem 96,2% de acurácia, também em textos de língua escrita. O LX-Tagger<sup>4</sup> é uma ferramenta que foi desenvolvida na

<sup>4</sup> <http://lxcenter.di.fc.ul.pt/tools/pt/conteudo/LXTagger.html#pq>

Universidade de Lisboa pelo NLX - Grupo de Fala e Linguagem Natural do Departamento de Informática. LX-Tagger é um etiquetador morfossintático como o Aelius, e etiqueta seus *corpora* de forma semelhante, utilizando o símbolo / (barra) após cada ocorrência e definindo sua devida etiqueta. Um ponto positivo do LX-Tagger, e que difere do Aelius, é que ele é capaz de reconhecer expressões, marcando cada unidade da mesma com um L e em seguida um número que representa a posição da palavra na expressão, como podemos ver no exemplo abaixo:

“de maneira a que” → de/LCJ1 maneira/LCJ2 a/LCJ3 que/LCJ4

O etiquetador com o qual estamos trabalhando, o Aelius, não etiqueta da mesma maneira que o LX-Tagger as expressões, ele reconhece cada palavra isolada, não como parte de um conjunto, do ponto de vista de expressões.

#### **4.3.1 Principais problemas encontrados na correção manual**

O Aelius foi programado para analisar textos de língua escrita; por isso, encontramos alguns problemas ao usá-lo para etiquetar textos de língua falada, como esperado. Na língua falada temos ruídos, hesitações e truncamentos que acabam gerando palavras desconhecidas pelo etiquetador, que precisa encontrar uma solução em seu dicionário para gerar uma etiqueta para essa “pseudo-palavra”; há algumas interjeições menos comuns e expressões que o anotador tem dificuldade de etiquetar; e, além disso, há gírias e palavras muito típicas e específicas de Porto Alegre, que acabam sendo etiquetadas erradamente, acreditamos que, por não constarem no dicionário do programa.

Outros problemas encontrados, que não se restringem à língua falada, mas também aparecem na língua escrita, são que o etiquetador não reconhece a estrutura de nome próprio composto, o que acaba gerando erro em alguns casos e, ademais, há palavras que são polissêmicas ou homônimas, e o Aelius por vezes acaba etiquetando de acordo com um dos possíveis significados, mas que não é o correto naquele contexto.

Como sabemos, algumas palavras podem apresentar mais de um significado, como *banco*, por exemplo, que pode designar uma instituição financeira, um lugar para sentar, ou ainda, um verbo na primeira pessoa do singular do presente do indicativo. Por esse motivo, o etiquetador está autorizado a designar etiquetas diferentes para uma mesma palavra, o que

pode gerar uma etiqueta correta ou não. Algumas vezes, diferentes ocorrências de uma mesma palavra são anotadas com etiquetas diferentes, em contextos muito parecidos.

E, por fim, há algumas etiquetas do Aelius, que segundo a lista de etiquetas não estariam erradas, como por exemplo, meses do ano e dias da semana marcados como nome próprio. Porém, com a finalidade de sofisticarmos o anotador, sugerimos que sejam usadas etiquetas de nome comum para esses termos, pois, segundo Bechara (2009: 113):

Substantivo comum é o que se aplica a um ou mais objetos particulares que reúnem características inerentes a dada classe: homem, mesa, livro, cachorro, lua, sol, fevereiro, segunda-feira, papa. Os cinco últimos exemplos patenteiam que há substantivos comuns que são nomes individualizados, não como os nomes próprios, mas pelo contexto extralinguístico e pelo nosso saber que nos diz que no contexto “natural” nosso só há uma lua, um sol, um mês de fevereiro e um só dia da semana segunda-feira e, no contexto “cultural”, só há um papa. Se forem escritos com maiúscula, deve-se o fato a pura convenção ortográfica, e não porque são nomes próprios.

Porém, neste trabalho proporemos soluções apenas para os erros que dizem respeito a características típicas de língua falada, como vemos na seção 4.2.1.2.

Separamos os erros em dois tipos: os erros de “confusão” de etiquetas e os erros por limitações do programa frente a dados típicos de língua falada. A seção 4.2.1.1 é uma sistematização dos erros de “confusão” de etiquetas, em que apresentaremos as ocorrências desses erros em seus contextos. A seção 4.2.1.2 é uma sistematização dos erros encontrados na anotação de língua falada especificamente.

#### **4.3.1.1 Erros de “confusão” de etiquetas**

Nesta seção, apresentamos os erros do Aelius ao etiquetar palavras com etiquetas já “conhecidas”, ou seja, palavras cuja classe morfológica é conhecida pelo programa. A seguir, ilustraremos esses problemas encontrados com exemplos que apareceram no *corpus* em seu contexto original. A ideia básica aqui é listar e sistematizar absolutamente *todos* os erros que o anotador cometeu ao analisar os trechos de língua falada do Varsul que anotou. Acreditamos que nossas anotações possam contribuir, dessa forma, para novas e melhores versões do etiquetador.

Na tabela 3, apresentaremos as 16 ocorrências de adjetivos que foram etiquetados pelo Aelius como nomes.

OCORRÊNCIA	CONTEXTO
<b>Arrumadinho</b>	... aquilo/DEM tudo/Q bem/ADV <i>arrumadinho</i> /N ./, aquele/D
<b>Baita</b>	... um/D-UM <i>baita</i> /N de/P um/D-UM salão/N
<b>Calma</b>	... uma/D-UM-F cidade/N muito/Q mais/ADV-R <i>calma</i> /N ./,
<b>Caretas</b>	... eles/PRO são/SR-P os/D-P <i>caretas</i> /N-P na/P+D-F realidade/N
<b>Culturais</b>	... <i>Porto</i> /NPR <i>Alegre</i> /NPR ./, aspectos/N-P <i>culturais</i> /N-P ./, o/D que
<b>Doidão</b>	... era/SR-D muito/Q <i>doidão</i> /N e/CONJ muito/Q <i>viajandão</i> /N ./.
<b>Durão</b>	... não/NEG pode/VB-P ser/SR-P <i>durão</i> /N demais/ADV-R ./, tem/TR-P
<b>Financeiro</b>	... e/CONJ o/D retorno/N <i>financeiro</i> /N também/ADV não/NEG era/SR-D
<b>Física</b>	... espécie/N de/P doença/N :/. <i>física</i> /N ./, psicológica/ADJ-F ./, espiritual
<b>Guri</b>	... quando/CONJS eu/PRO era/SR-D <i>guri</i> /N ./, quase/ADV não/NEG
<b>Histórica</b>	... <i>né</i> /NPR ?/. A/D-F parte/N <i>histórica</i> /N ./, mas/CONJ é/SR-P
<b>Miserável</b>	... não/NEG era/SR-D pobre/ADJ-G <i>miserável</i> /N como/CONJS hoje
<b>Passageira</b>	... pra/P mim/PRO uma/D-UM-F coisa/N <i>passageira</i> /N mesmo/ADJ ./.
<b>Sozinhos</b>	... então/ADV ./, nos/PRO saíamos/VB-D <i>sozinhos</i> /N-P ./, íamos/VB-D a
<b>Viajandão</b>	... era/SR-D muito/Q <i>doidão</i> /N e/CONJ muito/Q <i>viajandão</i> /N ./.
<b>Xarope</b>	... ficar/VB <i>meio</i> /N ridículo/ADJ ./, meio/ADJ <i>xarope</i> /N

**Tabela 3:** Adjetivo etiquetado como nome

Na tabela 4, apresentaremos os 12 adjetivos que foram etiquetados como nomes próprios.

OCORRÊNCIA	CONTEXTO
<b>Bastante</b>	... e/CONJ <i>Bastante</i> /NPR coisa/N ,/, a/D-F gente/N vai/VB-P morrer/VB
<b>Durão</b>	... mas/CONJ o/D pai/N era/SR-D <i>durão</i> /NPR ./ E/CONJ E/CONJ a/D-F
<b>Ecológica</b>	... um/D-UM pouquinho/N lá/ADV na/P+D-F <i>ecológica</i> /NPR ,/, <i>né</i> /NPR
<b>Eletrônico</b>	... entrou/VB-D ,/, que/C o/D <i>portão</i> /NPR <i>eletrônico</i> /NPR abriu/VB-D ,/,
<b>Garçom</b>	... já/ADV trabalhaste/VB-D de/P <i>garçom</i> /NPR também/ADV
<b>Grávida</b>	... veio/VB-D pro/P+D Brasil/NPR <i>grávida</i> /NPR do/P+D meu/PRO\$ avô
<b>Interiorzão</b>	... interior/N assim/ADV então/ADV é/SR-D <i>interiorzão</i> /NPR assim
<b>Mãezona</b>	... eu/PRO fui/SR-D bem/ADV <i>mãezona</i> /NPR deles/P+PRO-P ./.
<b>Monótono</b>	... achando/VB-G muito/Q <i>monótono</i> /NPR ir/VB com/P os/D-P pais/N-P
<b>Problemáticos</b>	... amigos/N-P ./ Inclusive/FP esses/D-P <i>problemáticos</i> /NPR-P ,/, que
<b>Tranqüilo</b>	... É/SR-P ,/, eu/PRO me/CL sinto/VB-P <i>tranqüilo</i> /NPR
<b>Tranqüilos</b>	... ,/, ficávamos/VB-D ali/ADV <i>tranqüilos</i> /N-P <i>né</i> /NPR ?/.

**Tabela 4:** Adjetivo etiquetado como nome próprio

Na tabela 5, está a única ocorrência de adjetivo etiquetado como numeral, e se trata de um caso muito específico, em que a forma do adjetivo é a mesma forma do numeral.

OCORRÊNCIA	CONTEXTO
<b>Sete</b>	... futebol/N <i>sete</i> /NUM e/CONJ a/D-F gente/N vê/VB-P assim/ADV

**Tabela 5:** Adjetivo etiquetado como numeral

Na tabela 6, temos as 14 ocorrências de adjetivos etiquetados verbos.



OCORRÊNCIA	CONTEXTO
<b>Brigão</b>	... ./, assim/ADV ./, <i>machista</i> /VB-P ./, <i>brigão</i> /VB-P ./, não/NEG sabe
<b>Chata</b>	... mas/CONJ deve/VB-P ser/SR bem/ADV <i>chata</i> /VB-AN-F ./, <i>né</i> /NPR
<b>Correto</b>	... tudo/Q bem/ADV certo/ADJ ./, bem/ADV <i>correto</i> /VB-P e/CONJ
<b>Dourado</b>	... é/SR-P um/D-UM sonho/N <i>dourado</i> /VB-AN ./, eu/PRO acho/VB-P
<b>Dourados</b>	... eu/PRO sou/SR-P dos/P+D-P anos/N-P <i>dourados</i> /VB-AN-P
<b>Doutorando</b>	... então/ADV ele/PRO é/SR-P <i>doutorando</i> /VB-G ./, dá/VB-P aula/N
<b>Escolar</b>	... num/P+D-UM grupo/N <i>escolar</i> /VB ali/ADV ./, estudei/VB-D
<b>Gentis</b>	... são/SR-P muito/Q <i>gentis</i> /VB-P ./, toda/Q-F vez/N
<b>Hospitalar</b>	... padrão/N ./, com/P administração/N <i>hospitalar</i> /VB ./, que/C
<b>Machista</b>	... troglodita/ADJ-G ./, assim/ADV ./, <i>machista</i> /VB-P ./, <i>brigão</i> /VB-P
<b>Povão</b>	... o/D <i>Lula</i> /N muito/Q <i>povão</i> /VB-P pra/P ser/SR um/D-UM presidente/N
<b>Preliminar</b>	... a/D-F parte/N <i>preliminar</i> /VB de/P de/P teor/N de/P parte/N
<b>Rudimentar</b>	... lógico/ADV que/C muito/Q <i>rudimentar</i> /VB e/CONJ nada/N
<b>Xis</b>	... num/P+D bom/ADJ raio/N <i>xis</i> /VB-P ./, pra/P ocupar/VB aquele/D

Tabela 6: Adjetivo etiquetado como verbo

Na tabela 7, estão esquematizadas as 6 ocorrências dos advérbios que foram etiquetados como adjetivos.

OCORRÊNCIA	CONTEXTO
<b>Bastante</b>	... o/D senhor/NPR viaja/VB-P <i>bastante</i> /ADJ-R-G ?./
<b>Diferente</b>	... o/D portoalegrense/N fala/VB-P <i>diferente</i> /ADJ-G do/P+
<b>Direto</b>	... queria/VB-D tocar/VB <i>direto</i> /ADJ ./, assim/ADV um/D-UM <i>pós</i>
<b>Embaixo</b>	... e/CONJ botavam/VB-P <i>embaixo</i> /ADJ daquela/D-F argola/N e/CONJ
<b>Mesmo</b>	... todo/Q mundo/N se/SE recolher/VB <i>mesmo</i> /ADJ ./, porque/CONJ
<b>Prontamente</b>	... ela/PRO <i>prontamente</i> /ADJ-G arrumou/VB-D um/D-UM quarto/N

Tabela 7: Advérbio etiquetado como adjetivo

Na tabela 8, temos as 3 ocorrências de advérbios que foram etiquetados como nomes.

OCORRÊNCIA	CONTEXTO
Bem	... <i>Bem</i> /N provável/ADJ-G ./, <i>né</i> /NPR ?/.
Direitinho	... tiraram/VB-D ./, cortaram/VB-D <i>direitinho</i> /N a/D-F bolsa/N
Meio	... ficar/VB <i>meio</i> /N ridículo/ADJ ./,

Tabela 8: Advérbio etiquetado como nome

Na tabela 9, temos os 4 casos de advérbios que foram etiquetados como nomes próprios.

OCORRÊNCIA	CONTEXTO
Ali	... já/ADV morei/VB-D na/P+D-F Olinda/NPR <i>ali</i> /NPR ./,
Aqui	... o/D Adilson/NPR ./, <i>amigão</i> /NPR nosso/PRO\$ <i>aqui</i> /NPR ./, <i>né</i> /NPR
Cedinho	... isso/DEM aí/ADV ./, ó/INTJ !/. <i>Cedinho</i> /NPR ./, ah/INTJ !/. Eu/PRO
Muitíssimo	... cresceu/VB-D !/. Cresceu/VB-D <i>muitíssimo</i> /NPR !/. Da/P+D-F altura

Tabela 9: Advérbio etiquetado como nome próprio

Na tabela 10, há as 3 ocorrências de advérbios que foram etiquetados como verbos.

OCORRÊNCIA	CONTEXTO
Acá	... lá/ADV e/CONJ acolá/ADV ./, não/NEG é/SR-P <i>acá</i> /VB
Allá	... não/NEG é/SR-P <i>acá</i> /VB nem/CONJ-NEG <i>allá</i> /VB !/.
Dai	... é/SR-P é/SR-P um/D-UM <i>dai</i> /VB-I um/D-UM um/D-UM dos/P+D-P

Tabela 10: Advérbio etiquetado como verbo

Na tabela 11, está a única ocorrência de artigo etiquetado como clítico.

OCORRÊNCIA	CONTEXTO
Os	... <i>engulo</i> /N <i>os</i> /CL “/QT <i>esses</i> /NPR “/QT e/CONJ <i>os</i> /CL “/QT <i>erres</i> /N-P

Tabela 11: Artigo etiquetado como clítico

Na tabela 12, temos a ocorrência do único caso de artigo etiquetado como preposição.

OCORRÊNCIA	CONTEXTO
A	... e/CONJ <b>a</b> /P <i>Nenê</i> /VB que/C o/CL digam/VB-SP

Tabela 12: Artigo etiquetado como preposição

Na tabela 13, há a o único artigo que foi etiquetado como verbo.

OCORRÊNCIA	CONTEXTO
O	... de/P repente/N ./, <b>o</b> /CL <i>guri</i> /VB-D ser/SR criado/VB-NA por/P

Tabela 13: Artigo etiquetado como verbo

Na tabela 14, temos o único caso de conjunção etiquetada como preposição.

OCORRÊNCIA	CONTEXTO
Salvo	...se/SE Deus/NPR quiser/VB-SR ./, <b>salvo</b> /PREP algum/Q problema/N

Tabela 14: Conjunção etiquetada como preposição

Na tabela 15, há as 2 ocorrências de interjeições etiquetadas como adjetivos.

OCORRÊNCIA	CONTEXTO
Poxa	... muito/Q agito/N ./, <i>né</i> /NPR ?/, <b>Poxa</b> /ADJ-F !/. Eu/PRO estou/VB-P
Pronto	... todos/Q-P realizados/VB-AN-P ./, <b>Pronto</b> /ADJ ./, estão/ET-P aí/ADV

Tabela 15: Interjeição etiquetada como adjetivo

Na tabela 16, estão sistematizados os 5 casos de interjeições etiquetadas como nomes.

OCORRÊNCIA	CONTEXTO
Ai	... interessante/ADJ-G ./, <b>Ai</b> /N ./, desculpe/VB-SP
Bah	... aquilo/DEM ali/ADV era/SR-D <b>bah</b> /N !/. Era/SR-D bem/ADV de/P
Barbaridade	... muito/Q bom/ADJ de/P cozinha/N ./, <b>barbaridade</b> /N !/.
Credo	... coisa/N de/P louco/N ./, <b>credo</b> /N ./.
Poxa	... Digo/VB-P :/. “/QT <b>Poxa</b> /N !/. “/QT E/CONJ não/NEG é/SR-P

Tabela 16: Interjeição etiquetada como nome

Na tabela 17, podemos ver as 9 ocorrências de interjeições etiquetadas como nomes próprios.

OCORRÊNCIA	CONTEXTO
<b>Ai</b>	... <b>Ai/NPR</b> ./, alegria/N vou/VB-P te/CL contar/VB a/D-F gente/N
<b>Bah</b>	... <b>Bah/NPR</b> !/. Vai/VB-P ser/SR horrível/ADJ-G
<b>Hãn</b>	... Pra/P onde/WADV é/SR-P que/C tu/PRO vais/VP-P ?/. <b>Hãn/NPR</b> ?/.
<b>Ô</b>	... <b>Ô/NPR</b> !/. Cláudio/NPR !/. É/SR-P mas/CONJ é/SR-P ./, estava/ET-D
<b>Ôpa</b>	... eu/PRO digo/VB-P ./: <b>Ôpa/NPR</b> tem/TR-P duas/NUM-F ruas/N-P
<b>Pô</b>	... <b>Pô/NPR</b> !/. Mandar/VB o-D cara/N embora/ADV porquê/WADV ?/.
<b>Ué</b>	... diz/VB-P assim/ADV ./: “/QT <b>Ué/NPR</b> !/. Porque/CONJ mãe/N
<b>Tchê</b>	... sério/ADJ ./, viu/VB-D ?/. <b>Tchê/NPR</b> ./: E/CONJ eu/PRO vejo/VB-D
<b>Xi</b>	... mudou/VB-P !/. <b>Xi/NPR</b> !/. Progrediu/VB-D !/. Quando/CONJS

Tabela 17: Interjeição etiquetada como nome próprio

Na tabela 18, estão listados os 4 casos de interjeições etiquetadas como verbos.

OCORRÊNCIA	CONTEXTO
<b>Chê</b>	... nosso/PRO\$ braço/N direito/ADJ aqui/ADV ./, <b>chê/VB</b> !/.
<b>Puxa</b>	... <b>Puxa/VB-I</b> !/. Mas/CONJ será/SR-R que/C eu/PRO ia/VB-D
<b>Tomara</b>	... ali/ADV ./: <b>Tomara/VB-RA</b> que/C seja/VB-SP um/D-UM ambiente/N
<b>Tchê</b>	... Vem/VB-P cá/ADV <b>tchê/VB</b> ./, aqui/ADV não/NEG vai/VB-P ./,

Tabela 18: Interjeição etiquetada como verbo

Na tabela 19 estão as 52 ocorrências de nomes etiquetados como adjetivos.

OCORRÊNCIA	CONTEXTO
<b>Alemão</b>	... colônia/N alemã/AD-F e/CONJ o/D <b>alemão/ADJ</b> é/SR-P muito/Q
<b>Aniversário</b>	... dito/VB-AN é/SR-P <b>aniversário/ADJ</b> de/P amigos/N-P ./.
<b>Área:</b>	... que/C a/D-F <b>área/ADJ-F</b> mais/ADV difícil/ADJ-G de/P
<b>Astral</b>	... maravilhosa/ADJ-F ./, né/NPR ?/. <b>Astral/ADJ-G</b> fantástico/ADJ
<b>Atrativo</b>	... ali/ADV tinha/TR-D um/D-UM <b>atrativo/ADJ</b> ./, acho/VB-P que/C é

<b>Branco</b>	... Porque/CONJ o/D <b>branco</b> /ADJ leva/VB-P doença/N ./, leva/VB-P
<b>Brinquedo</b>	... pensar/VB num/P+D-UM <b>brinquedo</b> /ADJ e/CONJ imaginar/VB um
<b>Brincadeiras</b>	... as/D-F-P <b>brincadeiras</b> /ADJ-F-P que/WPRO as/D-F-P crianças/N-P
<b>Cívica</b>	... que/C é/SR-D <b>Moral</b> /ADJ-G e/CONJ <b>Cívica</b> /ADJ-F pra/P você/PRO
<b>Colegial</b>	... depois/ADV ia/VB-D até/P fazer/VB o/D <b>colegial</b> /ADJ-G do/P+D
<b>Cólica</b>	... a/D-F criança/N tem/TR-P <b>cólica</b> /ADJ-F ./, chora/VB-P a/D-F
<b>Consultório</b>	... filhos/N-P ./, é/SR-P <b>consultório</b> /ADJ ./, é/SR-P no/P+D emprego/N
<b>Cotidiano:</b>	... problemas/N-P do/P+D <b>cotidiano</b> /ADJ do/P+D bairro/N
<b>Diesel</b>	... porque/CONJ se/SE gasta/VB-P muito/Q <b>diesel</b> /ADJ-G F/NPR Posto
<b>Enxoval</b>	... fazer/VB o/D <b>enxoval</b> /ADJ-G ./, né/NPR ?/.
<b>Escuro</b>	... gostou/VB-D de/P dormir/VB no/P+D <b>escuro</b> /ADJ diretamente/ADV
<b>Espanhol</b>	... países/N-P falam/VB-P <b>espanhol</b> /ADJ-G ./, né/NPR ?/.
<b>Francês</b>	... noção/N muito/Q rápida/ADJ-F de/P inglês/N ./, de/P <b>francês</b> /ADJ
<b>Futuro</b>	... vida/N ./ O/D <b>futuro</b> /ADJ a/D-F gente/N também/ADV não/NEG
<b>Geral</b>	... Vai/VB-P dar/VB uma/D-UM-F <b>geral</b> /ADJ-G em/P todo/Q o/D
<b>Ginástica</b>	... uma/D-UM-F <b>ginástica</b> /ADJ-F mais/ADV-R pra/P de/P acordo/N
<b>Gravatinha</b>	... bota/VB-P anel/N de <b>grau</b> /N ./, bota/N <b>gravatinha</b> /ADJ-F ./, anda
<b>Grego:</b>	... língua/N mais/ADV-R bonita/ADJ-F ./, é/SR-P o/D <b>grego</b> /ADJ ./, ou
<b>Helicóptero</b>	... comprar/VB um/D-UM <b>helicóptero</b> /ADJ ./, um/D-UM uma/D-UM-F
<b>Hidrogenástica</b>	... estou/VB-P fazendo/VB-G essa/D-F <b>hidrogenástica</b> /ADJ-F ;/. aprendi
<b>Informática</b>	... Bom/ADJ ./, a/D-F <b>informática</b> /NPR em/P si/PRO ./, é/SR-P o/D
<b>Inglês</b>	... países/N-P falam/VB-P em/P <b>inglês</b> /ADJ ?/.
<b>Largo</b>	... ali/ADV no/P+D <b>largo</b> /ADJ assim/ADV que/C a/N-F criança da
<b>Local</b>	... depois/ADV no/P+D <b>local</b> /ADJ-G é/SR-P a/D-F guia/N ./.
<b>Lojinha</b>	... vão/VB-P ali/ADV na/P+D-F <b>lojinha</b> /ADJ-F <b>umenoventaenove</b> /N
<b>Maconha</b>	... gosto/N horrível/ADJ-G ./, <b>maconha</b> /ADJ ./, não/NEG gostei/VB-D
<b>Marginal</b>	... existe/VB-P muito/Q <b>marginal</b> /ADJ-P ./, muita/Q-F bebedeira/N
<b>Meia</b>	... pode/VB-P ser/SR um/D-UM par/N de/P <b>meia</b> /ADJ-F ./, um/D-UM
<b>Militar</b>	... assim/ADV que/C tinha/TR-D muito/Q <b>militar</b> /ADJ-G da/P+D
<b>Moral</b>	... o/D que/WPRO que/C é/SR-D <b>Moral</b> /ADJ-G e/CONJ <b>Cívica</b> /ADJ-F
<b>Negro</b>	... porque/CONJ Tem/TR-P <b>negro</b> /ADJ com/P dezessete/NUM anos/N-P
<b>Palco</b>	... o/D <b>palco</b> /ADJ pra/P nós/PRO
<b>Persianas</b>	... ainda/ADV tem/TR-P aquelas/D-F-P <b>persianas</b> /ADJ-F-P né/NPR ?/.

<b>Pessoal</b>	... o/D <i>pessoal</i> /ADJ-G pode/VB-P dar/VB a/D entender/VB
<b>Português</b>	... falo/VB-P corretamente/ADV o/D <i>português</i> /ADJ porque/CONJ
<b>Quarta-feira</b>	... <i>segunda</i> /ADJ-F ou/CONJ <i>quarta-feira</i> /ADJ-F ./, amanhã/ADV
<b>Quartos</b>	... quartos/N-P ./, são/SR-P <i>quartos</i> /NPR pra/P com/P banheiros/N-P
<b>Rituais</b>	... Que/WD tipo/N de/P <i>rituais</i> /ADJ-G-P vocês/PRO fazem/VB-P ?/.
<b>Rodoviária</b>	... fui/SR-D lá/ADV na/P+D-F <i>rodoviária</i> /ADJ-F ./, e/CONJ tinha/TR-D
<b>Terno</b>	... por/P exemplo/N ./, um/D-UM <i>terno</i> /ADJ bege/N bastante/ADJ-R-G
<b>Ternos</b>	... tingiam/VB-D os/D-P <i>ternos</i> /ADJ-P ./ As/D-F-P os/D-P tecidos
<b>Umbanda</b>	... parou/VB-D de/P frequentar/VB a/D-F <i>umbanda</i> /ADJ-F F/NPR Não
<b>Vaga</b>	... ela/PRO precisou/VB-D esperar/VB <i>vaga</i> /ADJ-F em/P avião/N pra/P
<b>Vagas</b>	... abrisse/VB-SD <i>vagas</i> /ADJ-F-P pra/P fazer/VB essa/D-F prova
<b>Versus:</b>	... questão/N entre/P índios/N-P <i>versus</i> /ADJ-P garimpeiros/N-P ?/
<b>Visual</b>	... e/CONJ cuida/VB-P do/P+D <i>visual</i> /ADJ-G ./, <i>né</i> /NPR ?/ E/CONJ vai
<b>Vizinho</b>	... não/NEG sou/SR-P muito/Q de/P ir/VB visitar/VB <i>vizinho</i> /ADJ ./,

Tabela 19: Nome etiquetado como adjetivo

Na tabela 20, há as 6 ocorrências de nomes etiquetados como advérbios.

<b>OCORRÊNCIA</b>	<b>CONTEXTO</b>
<b>Boate</b>	... Era/SR-D era/SR-D <i>boate</i> /ADV ali/ADV ?/.
<b>Cerca</b>	... foram/VB-D fazendo/VB-G <i>cerca</i> /ADV ./ Quando/CONJS a/D-F
<b>Lanche</b>	... eu/PRO prefiro/VB-P fazer/VB <i>lanche</i> /ADV mesmo/ADJ ./, <i>né</i> /NPR
<b>Lixo</b>	... colocando/VB-G <i>lixo</i> /ADV ./, até/P a/D-F parte/N comunitária/ADJ-F
<b>Mamãe</b>	... <i>papai</i> /VB-I e <i>mamãe</i> /ADV não/NEG é/SR-P vida/N toda/Q-F
<b>Sacola</b>	... eu/PRO ia/VB-D com/P <i>sacola</i> /ADV fazer/VB compras/N-P de/P

Tabela 20: Nome etiquetado como advérbio

Na tabela 21, há a única ocorrência de nome etiquetado como clítico.

<b>OCORRÊNCIA</b>	<b>CONTEXTO</b>
<b>Futebol</b>	... inclusive/ADV eu/PRO joguei/VB-D <i>futebol</i> /CL ali/ADV ./, no/P+D

Tabela 21: Nome etiquetado como clítico

Na tabela 22, estão os 101 casos de nomes comuns etiquetados como nomes próprios.

OCORRÊNCIA	CONTEXTO
Agosto	... me/CL aposentei/VB-D em/P <i>julho</i> /NPR e/CONJ <i>agosto</i> /NPR já
Aluguel	... uma/D-UM-F casa/N de/P <i>aluguel</i> /NPR lá/ADV perto/ADV
Amigão	... o/D Adilson/NPR ./, <i>amigão</i> /NPR nosso/PRO\$ <i>aqui</i> /NPR ./, <i>né</i> /NPR
Ano	... também/ADV ./, o/D <i>Ano</i> /NPR Novo/ADJ também/ADV ./, quer
Ascensão	... essa/D-F prova/N de/P <i>ascensão</i> /NPR ./, não/NEG só/FP pra/P mim
Auê	... criança/N brincando/VB-G ./, é/SR-P aquele/D <i>auê</i> /NPR <i>né</i> /NPR ?/.
Avião	... sobe/VB-P escadinha/N de/P <i>avião</i> /NPR <i>tranquilo</i> /NPR ./.
Babá	... tem/TR-P uma/D-UM-F <i>babá</i> /NPR cuja/WPRO\$-F mãe/N mora
Bagunça	... é/SR-P que/C virava/VB-D <i>bagunça</i> /NPR isso/DEM aqui/ADV
Banho	... banhos/N-P de/P óleo/N nos/P+D-P soldados/N-P <i>banho</i> /NPR de/P
Barranco	... não/NEG tinha/TR-D um/D-UM <i>barranco</i> /NPR na/P+D-F estrada/N
Básico	... informática/N eu/PRO fiz/VB-D só/FP o/D <i>básico</i> /NPR ./, <i>né</i> /NPR ?/.
Bolo	... receita/N de/P quê/WPRO ?/. E/CONJ <i>Bolo</i> /NPR ./, pão/N
Caçula	... ./, o/D <i>caçula</i> /NPR teve/TR-D um/D-UM problema/N
Calçamento	... nós/PRO não/NEG tínhamos/TR-D <i>cans</i> /N-P <i>calçamento</i> /NPR
Câmara	... fiz/VB-D parte/N na/P+D-F <i>Câmara</i> /NPR ./, entrei/VB-D com/P
Camarão	... frutos/N-P do/D+P mar/N ./, peixe/N ./, <i>camarão</i> /NPR ./, lula/N
Caminhão	... motorista/N de/P <i>caminhão</i> /NPR ?/. F/NPR De/P <i>caminhão</i> /NPR ./,
Campus	... ficava/VB-D bem/ADV ali/ADV no/P+D <i>Campus</i> /NPR ./, <i>né</i> /NPR
Chaminé	... vejo/VB-P a/D-F <i>chaminé</i> /NPR ali/ADV da/D-F usina/N e/CONJ
Chimarrão	... tomavam/VB-D “/QT”)/( <i>chimarrão</i> /NPR ./, churrasco/N acho/VB-P
Comando	... na/P+D-F sala/N de/P <i>comando</i> /NPR botaram/VB-D
Condomínio	... assaltaram/VB-D o/D escritório/N do/P+D <i>condomínio</i> /NPR ?/.
Correção	... ./, aparelho/N de/P <i>correção</i> /NPR ./, <i>né</i> /NPR ?/. como/CONJS
Criaçãozinha	... enfim/ADV ./, tinha/TR-D sua/PRO\$ <i>criaçãozinha</i> /NPR <i>né</i> /NPR ?/.
Currículo	... fazia/VB-D parte/N do/P+D <i>currículo</i> /NPR ./, <i>né</i> /NPR ?/. O/D
Década	... falou/VB-D na/P+D-F <i>década</i> /NPR da/P+D-F evaneglização/N
Dezembro	... foi/SR-D marcado/VB-D dia/N oito/NUM de/P <i>dezembro</i> /NPR ./, o/D
Divertimentos	... e/CONJ <i>divertimentos</i> /NPR quaisquer/Q-G-P ./.
Domingo	... tempos/N-P vagos/ADJ-P ?/. <i>Domingo</i> /NPR ?/. Eu/PRO gosto/VB-P

<b>Doutor</b>	... que/WPRO o/D <i>doutor</i> /NPR vinha/VB-D chegando/VB-G do/P+D
<b>Especialização</b>	... um/D-UM curso/N de/P <i>especialização</i> /NPR ./, <i>né</i> /NPR ?/.
<b>Espírita</b>	... Jesus/NPR que/C o/D <i>espírita</i> /NPR prega/VB-P pra/P que/C
<b>Esposo</b>	... rua/N ./O/D meu/PRO\$ <i>esposo</i> /NPR sempre/ADV disse/VB-P assim
<b>Ésses</b>	... <i>engulo</i> /N os/CL “/QT <i>ésses</i> /NPR “/QT e/CONJ os/CL “/QT erres/N-P
<b>Evolução</b>	... é/SR-P uma/D-UM-F questão/N de/P <i>evolução</i> /NPR ./ “/QT Temos
<b>Folders</b>	... peguei/VB-D na/P+D-F ./ aqueles/D-P <i>folders</i> /NPR-P ./, uma
<b>Fricção</b>	... aqueles/D-P brinquedinhos de/P de/P “/QT <i>fricção</i> /NPR “/QT ./, e
<b>Gás</b>	... foi/SR-D o/D <i>gás</i> /NPR aquele/D e/CONJ aí/ADV por/P último/N
<b>Gaúcho</b>	... se/SE o/D <i>gaúcho</i> /NPR tiver/TR-SR que/C andar/VB
<b>Gaúchos</b>	... sotaque/N diferente/ADJ-G dos/P+D-P <i>gaúchos</i> /NPR aqui/ADV
<b>Ginásio</b>	... depois/ADV do/P+D <i>ginásio</i> /NPR ./, que/WPRO também/ADV ia
<b>Ginástica</b>	... idade/N ./, porque/CONJ daí/P+ADV a/P <i>ginástica</i> /NPR ./, (( “/QT
<b>Gírias</b>	... agora/ADV ./, <i>gírias</i> /NPR-P ./, isso/DEM aí/ADV então
<b>Hidroginástica</b>	... eu/PRO faço/VB-P minha/PRO\$ <i>hidroginástica</i> /NPR lá/ADV na
<b>Horário</b>	... Porque/CONJ o/D <i>horário</i> /NPR lá/ADV é/SR-P das/P+D-F-P três
<b>Início</b>	... estava/ET-D falando/VB-G ./, <i>início</i> /NPR de/P <i>novembro</i> /NPR
<b>Íterim</b>	... sabe/VP-P ?/. Nesse/P+S meu/PRO\$ <i>íterim</i> /NPR ./, aguardando
<b>Irmã</b>	... só/FP a/D-F minha/PRO\$ <i>irmã</i> /NPR e/CONJ meu/PRO\$ pai/N
<b>Julho</b>	... me/CL aposentei/VB-D em/P <i>julho</i> /NPR e/CONJ <i>agosto</i> /NPR já/ADV
<b>Lampião</b>	... eu/PRO sou/VB-P do/P+D tempo/N do/P+D <i>lampião</i> /NPR ./.
<b>Latim</b>	... da/P+D-F língua/N dentro/ADV do/P+D <i>latim</i> /NPR ./, <i>né</i> /NPR ?/.
<b>Líder</b>	... como/CONJS é/SR-P ?/. <i>Líder</i> /NPR do/P+D grupo/N ./, conseguiu
<b>Líderes</b>	... eram/SR-D os/D-P <i>líderes</i> /NPR :/. Alceu/NPR Collares/NPR
<b>Língua</b>	... ?/. F/NPR <i>Língua</i> /NPR portuguesa/ADJ-F ?/. E/CONJ
<b>Mano</b>	... casa/N de/P tango/N <i>Mano</i> /NPR a/P <i>Mano</i> /NPR ./, de/P vez/N em/P
<b>Março</b>	... vinha/VB-D embora/ADV <i>início</i> /NPR de/P <i>março</i> /NPR
<b>Médica</b>	... sorte/N ./, porque/CONJ <i>médica</i> /NPR aí/ADV mesmo/ADJ que/C eles
<b>Metrô</b>	... abrir/VB pra/P fazer/VB o/D <i>metrô</i> /NPR ./, <i>né</i> /NPR ?/. Não/NEG sei
<b>Natal</b>	... criou/VB-D o/D <i>primeiro</i> /ADJ <i>Natal</i> /NPR do/P+D funcionário/N
<b>Nenê</b>	... preocupar/VB com/P um/D-UM <i>nenê</i> /NPR pequeno/ADJ ./.
<b>Novembro</b>	... a/P partir/VB do/P+D <i>mês</i> /N-P de/P <i>novembro</i> /NPR
<b>Ouro</b>	... opinar/VB ./, porque/CONJ é/SR-P <i>Ouro</i> /NPR ./, deve/VB-P ser/SR



<b>Padre</b>	... ./, médico/N ./, engenheiro/N ./, advogado ./, <b>padre/NPR</b> ./, tem/TR-P
<b>Palavrão</b>	... o/D tal/ADJ-R-G de/P <b>palavrão/NPR</b> ./, em/P teatro/
<b>Paralelepípedo</b>	... aqui/ADV nessa/P+D-F no/P+D <b>paralelepípedo/NPR</b> da/P+D-F
<b>Paredão</b>	... agarrada/VB-AN-F no/P+D <b>paredão/NPR</b> do/P+D
<b>Paróquia</b>	... aqui/ADV na/P+D-F nossa/PRO\$ <b>paróquia/NPR</b> mesmo/ADJ
<b>Picolé</b>	... uns/D-UM-P carrinhos/N-P faziam/VB-D <b>picolé/NPR</b> ./ Isso/DEM
<b>Plantão</b>	... oito/NUM horas/N-P de/P <b>plantão/NPR</b> ./, digamos/VB-SP ./, oito
<b>Plástico</b>	... é/SR-P aquelas/DEM-F-P de/P <b>plástico/NPR</b> ./, até/FP
<b>Portão</b>	... entrou/VB-D ./, que/C o/D <b>portão/NPR</b> eletrônico/NPR abriu/VB-D ./,
<b>Português</b>	... o/D ensino/N de/P <b>português/NPR</b> agora/ADV ?./
<b>Princesa</b>	... a/D-F <b>princesa/NPR</b> boa/ADJ-F atendia/VB-D os/D-P pedidos
<b>Professor</b>	... depois/ADV que/C bate/VB-P que/C o/D <b>professor/NPR</b> entra/VB-P
<b>Quintas</b>	... ?/. <b>Terças/NPR-P</b> e/CONJ <b>quintas/NPR-P</b> ./, a/D-F aula/N de/P
<b>Rádio</b>	... trabalhei/VB-D em/P <b>rádio/NPR</b> durante/P dez/NUM anos/N-P ./,
<b>Réguas</b>	... laboratórios/N-P ./, tipo/N <b>réguas/NPR</b> ./, <b>quebracabeça/NPR</b> ./,
<b>Reitoria</b>	... sair/VB da/P+D <b>Reitoria/NPR</b> pra/P minha/PRO\$ função/N
<b>Revólver</b>	... isso/DEM com/P o/D <b>revólver/NPR</b> na/P+D-F mão/N ./, <b>né/NPR</b> ?./
<b>Rodoviária</b>	... sair/VB do/P+D da/P+D-F <b>rodoviária/NPR</b> e/CONJ
<b>Sábado</b>	... madrugada/N de/P <b>sábado/NPR</b> ainda/ADV e/CONJ
<b>Sanduíche</b>	... a/D-F mortadela/N do/P+D <b>sanduíche/NPR</b> (/ (riso/N f/N) /)
<b>Secretaria</b>	... a/D-F <b>secretaria/NPR</b> já/ADV não/NEG <b>limpa/ADJ-F</b> mais/ADV-R
<b>Seqüestro</b>	... assalto/N do/P+D o/D <b>seqüestro/NPR</b> do/P+D seu/PRO\$ marido/N
<b>Senhor</b>	... um/D-UM <b>senhor/NPR</b> vinha/VB-D vindo/VB-G da/P+D-F
<b>Televisão</b>	... aquilo/DEM foi/SR-D a/D-F <b>televisão/NPR</b> ./, <b>né/NPR</b> ?./
<b>Terças</b>	... tarde/N <b>né/NPR</b> ?/. <b>Terças/NPR-P</b> e/CONJ <b>quintas/NPR-P</b> ./, a/D-F
<b>Térmica</b>	... larguei/VB-D a/D-F <b>térmica/NPR</b> ./, botamos/VB-D ele/PRO dentro
<b>Terra</b>	... também/ADV conheci/VB-P ./ <b>Terra/NPR</b> fria/ADJ-F lá/ADV ./
<b>Térreo</b>	... mora/VB-P no/P+D <b>térreo/NPR</b> estava/ET-D tão/ADV-R apavorada
<b>Tia</b>	... primeiro/ADV a/D-F minha/PRO\$ <b>tia/NPR</b> ./, a/D-F irmã/N dele
<b>Troço</b>	... é/SR-P um/D-UM <b>troço/NPR</b> assim/ADV ./, a/D-F casa/N
<b>Túnel</b>	... tu/PRO <b>passas/N-P</b> em/P <b>túnel/NPR</b> ./, ainda/ADV pode/VB-P
<b>Turismo</b>	... investindo/VB-G mais/ADV-R no/P+D <b>Turismo/NPR</b> ./, <b>né/NPR</b> ?./
<b>TV</b>	... depois/ADV da/P+D-F <b>TV/NPR</b> ./, TV/NPR Tupi/NPR ./,

<b>Uísque</b>	... garrafinha/N de/P <i>uísque</i> /NPR ./, essas/D-F-P de/P bolso/N
<b>Utensílios</b>	... fogão/N ./, de/P panela/N ./, <i>utensílios</i> /NPR-P domésticos/ADJ-P ./,
<b>Vídeo</b>	... num/P+D-UM programa/N de/P <i>vídeo</i> /NPR de/P televisão/N do/P+D
<b>Vó</b>	... a/D-F <i>vó</i> /NPR nunca/ADV-NEG deixou/VB-D a/D-F gente/N sair/VB
<b>WC</b>	... onde/WADV era/SR-D o/D <i>WC</i> /NPR eu/PRO não/NEG sabia/VB-D

**Tabela 22:** Nome etiquetado como nome próprio

Na tabela 23, estão as 9 ocorrências de nomes próprios etiquetados como adjetivos.

<b>OCORRÊNCIA</b>	<b>CONTEXTO</b>
<b>Americanas</b>	... quebraram/VB-D as/D-F-P Lojas/N-P <i>Americanas</i> /ADJ-F-P e/CONJ
<b>Argentina</b>	... da/P+D <i>Argentina</i> /ADJ-F e/CONJ do/P+D Uruguai/NPR
<b>Camaquã</b>	... voltei/VB-D pro/P+D bairro/N <i>Camaquã</i> /ADJ-F ./, depois/ADV
<b>Campinas</b>	... conheço/VB-P bem/ADV <i>Campinas</i> /ADJ-F-P ./ <i>Campinas</i> /N-P é/SR
<b>Copacabana</b>	... <i>Ipanema</i> /ADJ-S-F ou <i>Copacabana</i> /ADJ-F no/P+D Rio/NPR de/P
<b>Internacional</b>	... jogar/VB no/P+D <i>Internacional</i> /ADJ-G um/D-UM bom/ADJ tempo
<b>Ipanema</b>	... <i>Ipanema</i> /ADJ-S-F ou <i>Copacabana</i> /ADJ-F no/P+D Rio/NPR de/P
<b>Olímpico</b>	... dentro/ADV do/P+D <i>Olímpico</i> /ADJ ./, lá/ADV ./, num/P jogo/N
<b>Rafaela</b>	... quando/CONJS a/D-F <i>Rafaela</i> /ADJ-F nasceu/VB-D ./, eu/PRO

**Tabela 23:** Nome próprio etiquetado como adjetivo

Na tabela 24, estão os 25 nomes próprios etiquetados como nomes comuns.

OCORRÊNCIA	CONTEXTO
<b>Amazônia</b>	... vou/VB-P ter/TR ./ E/CONJ na/P+D-F <i>Amazônia</i> /N ?/. F/NPR Então
<b>Azaléia</b>	... trabalha/VB-P na/P+D-F <i>Azaléia</i> /N ./, e/CONJ
<b>Bagé</b>	... ter/TR que/C estar/ET em/P <i>Bagé</i> /N hem/INTJ !/.
<b>Brizola</b>	... e/CONJ o/D <i>Brizola</i> /N não/NEG lhe/CL conquistou/VB-D ?/.
<b>Cacalo</b>	... Então/ADV o/D <i>Cacalo</i> /N é/SR-P gremista/ADJ-G ./, né/NPR ?/.
<b>Campinas</b>	... conheço/VB-P bem/ADV <i>Campinas</i> /ADJ-F-P ./ <i>Campinas</i> /N-P é/SR
<b>Canoas</b>	... de/P morar/VB em/P <i>Canoas</i> /N-P ou/CONJ em/P Paris/NPR
<b>Colômbia</b>	... num/P+D-UM país/N como/CONJS a/D-F <i>Colômbia</i> /N ?/.
<b>Fortaleza</b>	... de/P albergue/N ./, Porto/NPR Seguro/NPR ./, <i>Fortaleza</i> /N ./, não
<b>Gasômetro</b>	... da/P+D-F <i>usina</i> /VB-AN-F aqui/ ADV do/P+D <i>Gasômetro</i> /N
<b>Getúlio</b>	... a/D-F morte/N do/P+D <i>Getúlio</i> /N acho/VB-P que/C dia/N vinte/NUM
<b>Grêmio</b>	... <i>Olimpico</i> /ADJ ./, lá/ADV ./, num/P jogo/N do/P+D <i>Grêmio</i> /N
<b>Guaíba</b>	... se/SE viu/VB-D o/D <i>Guaíba</i> /N botou/VB-D tudo/Q pra/P fora/ADV
<b>Hecléa</b>	... Então/ADV a/D-F <i>Hecléa</i> /N junto/ADV com/P Beatriz/NPR
<b>Lula</b>	... o/D <i>Lula</i> /N eu/PRO acho/VB-P que/C ia/VB-D fazer/VB perder/VB
<b>Montevideo</b>	... seguidamente/ADV vou/VB-P a/P <i>Montevideo</i> /N ./, e/CONJ a/D-F
<b>Olívio</b>	... foi/SR-D o/D <i>Olívio</i> /N nos/P+D-P quatro/NUM anos/N-P
<b>Pelotas</b>	... não/NEG tinha/TR-D em/P <i>Pelotas</i> /N naquela/P+D-F ocasião/N
<b>Rafaela</b>	... com/P a/D-F <i>Rafaela</i> /NPR ./, a/D-F <i>Rafaela</i> /N vai/VB-P comigo
<b>Redenção</b>	... o/D <i>Marinha</i> /NPR do/P+D <i>Brasil</i> /NPR ou/CONJ o/D <i>parque</i> /N da/P+D-F <i>Redenção</i> /N
<b>Romário</b>	... já/ADV chega/VB-P o/D <i>Romário</i> /N na/P+D <i>televisão</i> /N agora/ADV
<b>Rondônia</b>	... minha/PRO\$ ida/N a/P <i>Rondônia</i> /N ./ <i>Passei</i> /VB-D seis/NUM
<b>Rose</b>	... aí/ADV a/D-F <i>Rose</i> /N me/CL ensinou/VB-D a/P fazer/VB
<b>Sertório</b>	... tem/TR-P ali/ADV <i>embaixo</i> /ADJ ./, ali/ADV na/P+D-F <i>Sertório</i> /N
<b>UFRGS</b>	... e/CONJ a/D-F <i>UFRGS</i> /N também/ADV ./, que/WPRO eu/PRO

Tabela 24: Nome próprio etiquetado como nome comum

Na tabela 25, está o único caso de nome etiquetado como pronomes demonstrativos.

OCORRÊNCIA	CONTEXTO
<b>Esses</b>	... fala/VB-P com/P todos/Q-P os/D-P <i>esses/D-P</i> e/CONJ erres/N-P ./,

Tabela 25: Nome etiquetado como pronome demonstrativo

Na tabela 26, há a única ocorrência de nome etiquetado como pronomes interrogativos.

OCORRÊNCIA	CONTEXTO
<b>Bilboquê</b>	... Conhece/VB-P <i>bilboquê/VB ?/</i> . <i>Bilboquê/WD</i> Ah/INTJ !/. Tudo/Q

Tabela 26: Nome etiquetado como pronome interrogativo

Na tabela 27, temos os 72 casos de nomes que foram etiquetados como verbos.

OCORRÊNCIA	CONTEXTO
<b>Açúcar</b>	... ela/PRO bota/VB-P <i>açúcar/VB</i> fora/ADV do/P+D
<b>Adulto</b>	... <i>né/NPR ?/</i> . Só/FP <i>adulto/VB-P</i> em/P casa/N e/CONJ ele/PRO
<b>Aluna</b>	... fui/SR-D <i>aluna/VB-AN-F</i> ali/ADV ./, mas/CONJ não/NEG
<b>Bandidos</b>	... <i>né/NPR ?/</i> . Tem/TR-P <i>bandidos/VB-AN-P</i> lá/ADV ./, pessoas/N-P que
<b>Bilboquê</b>	... jogava/VB-D <i>bilboquê/VB</i> ./, Conhece/VB-P <i>bilboquê/VB ?/</i> .
<b>Bobagem</b>	... vai/VB-P afundar/VB ./, tudo/Q <i>bobagem/VB-D</i> ./, a/D-F água/N nem
<b>Briga</b>	... sempre/ADV dá/VB-P problema/N de/P <i>briga/VB-P</i> ./, de/P morte/N
<b>Brinquedo</b>	... mais/ADV-R em/P fazer/VB <i>brinquedo/VB-AN</i> ./, Pelo/P+D que
<b>Buraco</b>	... num/P+D-UM buraco/N assim/ADV e/CONJ entrou/VB-D <i>buraco/VB-G</i> adentro/ADV
<b>Calçados</b>	... fui/SR-D ser/SR entregador/N de/P <i>calçados/VB-AN-P</i> na/P+D-F
<b>Camarão</b>	... depois/ADV <i>pegase/ADV</i> o/D <i>camarão/VB-R</i> pronto/ADV ./, a/D
<b>Caminha</b>	... na/P+D-F <i>caminha/VB-P</i> do/P+D lado/N ./, daí/P+ADV a/D
<b>Caminhada</b>	... estão/ET-P numa/P+D-UM-F <i>caminhada/VB-AN-F</i> assim/ADV de/P
<b>Canis</b>	... tu/PRO não/NEG <i>escutas/N-P canis/VB-P</i> ./, têm/TR-P muito/Q
<b>Canivete</b>	... podia/VB-D chover/VB ./, cair/VB <i>canivete/VB-P</i> ./, lá/ADV
<b>Carioca</b>	... lá/ADV o/D <i>carioca/VB-D</i> ele/PRO também/ADV não/NEG é/SR-P
<b>Cartaz</b>	... parede/N e/CONJ cheia/ADJ-F de/P <i>cartaz/VB</i> e/CONJ

<b>Cercado</b>	... e/CONJ ficou/VB-D o/D <b>cercado/VB-AN</b> ali/ADV ./, então/ADV o/D
<b>Chamado</b>	... voz/N dele/P+PRO ./, <b>né/NPR</b> ?/. Pelo/P+D <b>chamado/VB-NA</b> dele
<b>Chimarrão</b>	... olha/VB-P <b>chimarrão/VB-R</b> a/D-F gente/N fazia/VB-D
<b>Conteúdo</b>	... tu/PRO <i>organizas/N-P um/D</i> <b>conteúdo/VB-AN</b> ./, mas/CONJ assim
<b>Contrapartida</b>	... assim/ADV ./, em/P <b>contrapartida/VB-AN-F</b> ./, estamos/VB-P
<b>Correria</b>	... então/ADV foi/SR-D aquela/D-F <b>correria/VB-R</b> ./,
<b>Criançada</b>	... assim/ADV que/C a/D-F <b>criançada/VB-AN-F</b> <i>brincava/N</i> ./ . Então
<b>Crochê</b>	... ./, ou/COM fazendo/VB-G <b>tricô/VB</b> e/CONJ <b>crochê/VB</b> ./, <b>né/NPR</b> ?./.
<b>Dentista</b>	... ./, de/P ser/SR <b>dentista/VB-AN-F</b> e/CONJ vendo/VB-P hoje/ADV a
<b>Desespero</b>	... chorando/VB-G de/P <b>desespero/VB-P</b> ./, então/ADV eu/PRO me/CL
<b>DMAE</b>	... por/P exemplo/N ./, <b>DMAE/VB-I</b> ./, CRT/NPR ./, a/D-F Prefeitura/N
<b>Emprego</b>	... mais/ADV-R <i>mãodeobra/N do/P+D</i> que/WPRO <b>emprego/VB-P</b> não
<b>Ensino</b>	... saúde/N precária/ADJ-F ./, <b>ensino/VB-P</b> nem/CONJ-NEG se/SE fala
<b>Enxurrada</b>	... deu/VB-D aquela/D-F <b>enxurrada/VB-AN-F</b> e/CONJ depois/ADV
<b>Erre</b>	... falava/VB-D aquilo/D puxando/VB-G no/P+D <b>erre/VB-SP</b> ./, sabe
<b>Folga</b>	... poucas/Q-F-P horas/N de/P <b>folga/VB-P</b> ./, <b>né/NPR</b> ?/.
<b>Franquia</b>	... lá/ADV <b>né/NPR</b> ?/. <b>Franquia/VB-D</b> disso/P+DEM ./, <b>franquia/VB-D</b>
<b>Fuligem</b>	... ainda/ADV derramava/VB-D <b>fuligem/VB-P</b> pra/P toda/Q-F
<b>Greve</b>	... eu/PRO vou/VB-P fazer/VB <b>greve/VB-P</b> ./, <b>né/NPR</b> ?/.
<b>Guri</b>	... de/P repente/N ./, o/CL <b>guri/VB-D</b> ser/SR criado/VB-NA por/P
<b>Gurizada</b>	... mas/CONJ era/SR-D <b>gurizada/VB-AN-F</b> mesmo/FP ./,
<b>Inseticida</b>	... porque/CONJS botam/VB-P muito/Q <b>inseticida/VB-AN-F</b> ./, muita
<b>Lampião</b>	... tempo/N do/P+D <i>lampião/NPR</i> ./ . <b>Lampião/VB-D</b> aquele/D como
<b>Maconha</b>	... amigo/N com/P portando/VB-G <b>maconha/VB-P</b> ./ . Estivemos/ET-D
<b>Mestrado</b>	... fazer/VB um/D-UM <i>mestrado/N</i> ./, não/NEG ./, <b>mestrado/VB-AN</b> por
<b>Nenê</b>	... e/CONJ a/P <b>Nenê/VB</b> que/C o/CL digam/VB-SP
<b>Novela</b>	... ah/INTJ ./, <b>novela/VB-P</b> eu/PRO assisto/VB-P essa/D-F depois/ADV
<b>Padaria</b>	... bom/ADJ morar/VB aqui/ADV ./, <b>padaria/VB-R</b> ./, mercearia/N
<b>Papai</b>	... <b>papai/VB-I</b> e <i>mamãe/ADV</i> não/NEG é/SR-P vida/N toda/Q-F
<b>Parada</b>	... chega/VB-P na/P+D-F <b>parada/VB-AN-F</b> assim/ADV ./, se/CONJS
<b>Passa</b>	... eu/PRO queria/VB-D comprar/VB <b>passa/VB-P</b> de/P uva/N e/CONJ
<b>Pijama</b>	... me/CL <i>animo/VB</i> só/FP fazendo/VB-G <b>pijama/VB-P</b> assim/ADV
<b>Pimentão</b>	... com/P tomate/N ./, cebola/N ./, <b>pimentão/VB-P</b> uma/D-UM-F camada

<b>Pós</b>	... tocar/VB <i>direto</i> /ADJ ./, assim/ADV um/D-UM <i>pós</i> /VB-D <i>né</i> /NPR ?/.
<b>Posto</b>	... muito/Q <i>diesel</i> /ADJ-G F/NPR <i>Posto</i> /VB-AN de/P gasolina/N ./,
<b>Quadra</b>	... casa/N ./, na/P+D-F mesma/ADJ-F <i>quadra</i> /VB-P da/P+D-F minha
<b>Rapazeada</b>	... hoje/ADV ./, será/VB-R <i>rapazeada</i> /VB-AN-F de/P amanhã/ADV ./,
<b>Revólver</b>	... aqui/ADV mas/CONJ com/P <i>revólver</i> /VB ali/ADV ./, ó/INTJ !/.
<b>Robôs</b>	... computação/N ./, <i>robôs</i> /VB-D é/SR-P ./, enfim/ADV ./, toda
<b>Saia</b>	... ./, e/CONJ <i>tule</i> /VB-SP ./, <i>né</i> /NPR ?/.
<b>Saia</b>	<i>Saia</i> /VB-SP de/P cetim/N com/P
<b>Segunda-feira</b>	... telefonou/VB-D pra/P casa/N <i>segundafeira</i> /VB-RA eu/PRO não/NEG
<b>Ser</b>	... é/SR-P próprio/ADJ do/P+D <i>ser</i> /SR humano/ADJ ./.
<b>Show</b>	... ./, fiz/VB-D <i>show</i> /VB na/P+D-F televisão/N ./, quatro/NUM anos/N-P
<b>Supermercado</b>	... um/D-UM pouco/N ./, foi/SR-D <i>supermercado</i> /VB-NA ./.
<b>Talco</b>	... se/SE enfeitava/VB-D ./, botava/VB-D <i>talco</i> /VB-G e/CONJ coisa/N
<b>Tapa</b>	... um/D-UM <i>tapa</i> /VB-P na/P+D-F <i>bunda</i> /N ./, daí/P+ADV
<b>Tecido</b>	... ./, quando/CONJS o/D <i>tecido</i> /VB-AN ficasse/VB-SD um/D-UM pouco
<b>Teclado</b>	... eu/PRO toco/VB-P <i>teclado</i> /VB-AN ./, mas/CONJ a/P órgão/N eu/PRO
<b>Tricô</b>	... ./, ou/COM fazendo/VB-G <i>tricô</i> /VB e/CONJ <i>crochê</i> /VB ./, <i>né</i> /NPR ?/.
<b>Tule</b>	... ./, e/CONJ <i>tule</i> /VB-SP ./, <i>né</i> /NPR ?/.
<b>Tule</b>	<i>Saia</i> /VB-SP de/P cetim/N com/P
<b>Turista</b>	... assim/ADV de/P repente/N como/CONJS <i>turista</i> /VB-P aquele/D
<b>Usina</b>	... lembro/VB-P da/P+D-F <i>usina</i> /VB-AN-F aqui/ ADV do/P+D
<b>Vestibular</b>	... quando/CONJS eu/PRO fiz/VB-D <i>vestibular</i> /VB pra/P faculdade/N
<b>Vestidos</b>	... vezes/N-P mais/ADV-R <i>vestidos</i> /VB-AN-P que/C essa/D-F menina/N
<b>Volta</b>	... uma/D-UM-F área/N tão/ADV-R grande/ADJ-G na/P+D-F <i>volta</i> /VB-P

Tabela 27: Nome etiquetado como verbo

Na tabela 28, temos o único caso de nome etiquetado como verbo + clítico.

<b>OCORRÊNCIA</b>	<b>CONTEXTO</b>
<b>Vídeogame</b>	... na/P+D-F frente/N do/P+D <i>videogame</i> /VB-SP+CL ./, e/CONJ só/FP

Tabela 28: Nome etiquetado como verbo + clítico

Na tabela 29, temos ocorrência de nome plural terminado em -S (cuja terminação também vale para o singular), etiquetado como nome plural.

OCORRÊNCIA	CONTEXTO
Ônibus	... passam/VB-P quatro/NUM <i>ônibus</i> /N quatro/NUM ou/CONJ cinco

Tabela 29: Nome plural terminado em -S etiquetado como nome singular

Na tabela 30, temos 2 ocorrências de nomes no singular terminados em -S etiquetados como nomes no plural.

OCORRÊNCIA	CONTEXTO
Mês	... a/P partir/VB do/P+D <i>mês</i> /N-P de/P <i>novembro</i> /NPR
Pós	... é/SR-P conseguir/VB fazer/VB um/D-UM <i>pós</i> /N-P ./.. Quero/VB-P

Tabela 30: Nome singular terminado em -S etiquetado como nome plural

Na tabela 31, temos a ocorrência de um numeral etiquetado como adjetivo.

OCORRÊNCIA	CONTEXTO
Primeiro	... criou/VB-D o/D <i>primeiro</i> /ADJ <i>Natal</i> /NPR do/P+D funcionário

Tabela 31: Numeral etiquetado como adjetivo

Na tabela 32, podemos ver a ocorrência de numeral etiquetado como clítico.

OCORRÊNCIA	CONTEXTO
Um	... daí/ADV cada/Q-G <i>um</i> /CL tinha/TR-D as/D-F-P suas/PRO\$

Tabela 32: Numeral etiquetado como clítico

Na tabela 33, há a ocorrência de um numeral etiquetado como pronome definido.

OCORRÊNCIA	CONTEXTO
Um	... tu/PRO <i>organizas</i> /N-P <i>um</i> /D <i>conteúdo</i> /VB-AN ./, mas/CONJ assim

Tabela 33: Numeral etiquetado como pronome definido

Na tabela 34, há o único caso de pronome demonstrativo etiquetado como nome próprio.

OCORRÊNCIA	CONTEXTO
Isso	... até/P <i>isso</i> /NPR ./, eu/PRO era/SR-D tão/ADV crente/ADJ-G

**Tabela 34:** Pronome demonstrativo etiquetado como nome próprio

Na tabela 35, há as 3 ocorrências de pronomes de tratamento etiquetados como nomes próprios.

OCORRÊNCIA	CONTEXTO
Dona	... aí/ADV ./, um/UM dia/N a/P <i>Dona</i> /NPR Alzira/NPR
Senhora	... quando/CONJS a/D-F <i>senhora</i> /NPR era/SR-D pequena/ADJ-F
Seu	... Olha/VB-P ./, <i>Seu</i> /NPR Eduardo/NPR ./, <i>Seu</i> /PRO\$ Aron/NPR esteve/ET-D lhe/PRO procurando/VB-G aí/ADV ./.

**Tabela 35:** Pronome de tratamento etiquetado como nome próprio

Na tabela 36, temos a ocorrência de pronome de tratamento etiquetado como pronome possessivo.

OCORRÊNCIA	CONTEXTO
Seu	... Olha/VB-P ./, <i>Seu</i> /NPR Eduardo/NPR ./, <i>Seu</i> /PRO\$ Aron/NPR esteve/ET-D lhe/PRO procurando/VB-G aí/ADV ./.

**Tabela 36:** Pronome de tratamento etiquetado como pronome possessivo

Na tabela 37, temos o único caso de pronome pessoal etiquetado como nome plural.

OCORRÊNCIA	CONTEXTO
Nós	... cinquenta/NUM e/CONJ seis/NUM <i>nós</i> /N-P casamos/VB-D

**Tabela 37:** Pronome pessoal etiquetado como nome plural

Na tabela 38, temos 6 ocorrências de pronomes pessoais etiquetados como nomes próprios.



OCORRÊNCIA	CONTEXTO
<b>Ela</b>	... <i>Ela/NPR</i> está/ET-P com/P um/D-UM problema/N pulmonar/ADJ-G ./,
<b>Ele</b>	... é/SR-P <i>Ele/NPR</i> está/ET-P acostumado/VB-AN com/P
<b>Eles</b>	... <i>Eles/NPR</i> moravam/VB-D na/P+D-F praia/N ./,
<b>Eu</b>	... garimpo/N de/P ouro/N <i>Eu/NPR</i> ./, um/D-UM amigo/N
<b>Nós</b>	... <i>Nós/NPR</i> brincávamos/VB-D assim/ADV ./, brigávamos/VB-D
<b>Tu</b>	... agora/ADV <i>Tu/NPR</i> saías/NPR-P limpinha/ADJ-F ./, arrumadinha

**Tabela 38:** Pronome pessoal etiquetado como nome próprio

Na tabela 39, temos a única ocorrência de pronome pessoal etiquetado como preposição + demonstrativo.

OCORRÊNCIA	CONTEXTO
<b>Nos</b>	... não/NEG <i>nos/P+D-P</i> <i>desestruturamos/N-P</i> ?/. Porque/CONJS a/D-F

**Tabela 39:** Pronome pessoal etiquetado como preposição + demonstrativo

Na tabela 40, há as 17 ocorrências de verbos etiquetados como adjetivos.

OCORRÊNCIA	CONTEXTO
Anulo	... há/HV-P muito/Q tempo/N eu/PRO <i>anulo</i> /ADJ ./, mas/CONJ eu/PRO
Baixa	... ponto/N aí/ADV o/CL ele/PRO <i>baixa</i> /ADJ-F ./, você/PRO se/SE
Boto	... pra/P casa/N ./, almoço/N ./, <i>boto</i> /ADJ a/D-F minha/PRO\$ roupa/N
Cola	... rasgadinha/ADJ-F ./, a/D-F gente/N <i>cola</i> /ADJ-F ./, bota/VB-P um
Curto	... assim/ADV ./, mas/CONJ até/P <i>curto</i> /ADJ ./, né/NPR ?/.
Encontras	... ./ tu/PRO <i>encontras</i> /ADJ-F-P casais/N-P lá/ADV ./, com/P
Estranha	... a/D-F comida/N a/D-F gente/N <i>estranha</i> /ADJ-F um/D-UM pouco/N
Estranhas	... então/ADV tu/PRO não/NEG <i>estranhas</i> /ADJ-F-P nada/Q-NEG ./,
Gostaria	... que/C a/D-F gente/N <i>gostaria</i> /ADJ-F ./, até/FP em/P termos/N-P de/P
Larga	... aí/ADV o/D cara/N <i>larga</i> /ADJ-F uma/D-UM-F música/N lá/ADV
Limpa	... a/D-F <i>secretaria</i> /NPR já/ADV não/NEG <i>limpa</i> /ADJ-F mais/ADV-R
Limpo	... né/NPR ?/. Eu/PRO <i>limpo</i> /ADJ caixa/N de/P gordura/N pra/P ela/PRO
Penduro	... máquina/N ./, lavo/VB-P ./, <i>penduro</i> /ADJ ./, depois/ADV vou/VB-P
Picas	... assim/ADV ./ tu/PRO <i>picas</i> /ADJ-F-P cebola/N ./, alho/N ./,
Precisa	... hora/N que/C a/D-F gente/N <i>precisa</i> /ADJ-F mesmo/ADJ
Provoca	... aqui/ADV ?/. F/NPR <i>Provoca</i> /ADJ-F problema/N de/P assalto/N
Transa	... jovem/N tal/ADJ-R-G ./, já/ADV <i>transa</i> /ADJ-F negócio/N de/P droga

Tabela 40: Verbo etiquetado como adjetivo

Na tabela 41, estão as 3 ocorrências de verbos etiquetados como advérbios.

OCORRÊNCIA	CONTEXTO
Existe	... aí/ADV ./ E/CONJ <i>existe</i> /ADV algum/Q outro/ADJ tipo/N de/P festa
Junto	... tem/TR-P que/C fazer/VB ./, né/NPR ?/. <i>Junto</i> /ADV o/N lixo/N ./ Tá
Reage	... casamento/N ./, a/D-F mulher/N <i>reage</i> /ADV está/ET-P ./, a/D-F

Tabela 41: Verbo etiquetado como advérbio

Na tabela 42, estão os 14 casos de verbos etiquetados como nomes.

OCORRÊNCIA	CONTEXTO
Acordo	... então/ADV eu/PRO me/CL <b>acordo</b> /N e/CONJ fico/VB-P ali/ADV de
Adoro	... eu/PRO <b>adoro</b> /N ./, estou/VB-P achando/VB-G maravilhoso
Almoço	... pra/P casa/N ./, <b>almoço</b> /N ./, boto/ADJ a/D-F minha/PRO\$ roupa/N
Bota	... amor/N de/P Deus/NPR !/. <b>Bota</b> /N sério/ADJ mesmo/ADJ ./.
Brincava	... assim/ADV que/C a/D-F <b>criançada</b> /VB-AN-F <b>brincava</b> /N ./.
Engulo	... <b>engulo</b> /N os/CL “/QT <b>esses</b> /NPR “/QT e/CONJ os/CL “/QT erres/N-P
Estimulo	... até/FP já/ADV <b>estimulo</b> /N isso/DEM ./, eu/PRO gosto/VB-P
Fala	... o/D <b>padre</b> /NPR sempre/ADV <b>fa</b> /VB <b>fala</b> /N que/WPRO não/NEG
Gosto	... muito/Q <b>gosto</b> /N muito/Q do/P+D lar/N ./, fico/VB-P
Jantar	... se/SE preparar/VB pra/P <b>jantar</b> /N e/CONJ coisa/N
Passeio	... ./, eu/PRO <b>passeio</b> /N bastante/ADJ-G ./, e/CONJ agora/ADV
Trabalho	... eu/PRO <b>trabalho</b> /N em/P outra/OUTRO-F cidade/N ./,
Tranca	... me/CL <b>tranca</b> /N muito/Q de/P viajar/VB
Volta	... sacola/N vai/VB-P vazia/ADJ-F ./.
	Na/P+D-F volta/N ./, <b>volta</b> /N cheia/

Tabela 42: Verbo etiquetado como nome

Na tabela 43, há as 3 ocorrências de verbos etiquetados como nomes próprios.

OCORRÊNCIA	CONTEXTO
É	... <b>É</b> /NPR ./, os portugueses/N-P já/ADV estão/VB-P também/ADV
Reúne	... a/D-F gente/N se/SE <b>reúne</b> /NPR numa/P+D-F sala/N de/P
Saiás	... que/C tu/PRO <b>saiás</b> /NPR-P ali/ADV naquela/P+D-F época

Tabela 43: Verbo etiquetado como nome próprio

Na tabela 44, há a única ocorrência de verbo conjugado etiquetado como verbo infinitivo.

OCORRÊNCIA	CONTEXTO
Vê	... nem/CONJ-NEG <b>vê</b> /VB as/D-F-P pessoas/N-P de/P repente/N vê

Tabela 44: Verbo conjugado etiquetado como verbo infinitivo

Na tabela 45, há os dois casos de verbos conjugados etiquetados como verbo *SER* no infinitivo.

OCORRÊNCIA	CONTEXTO
<b>Entrares</b>	... então/ADV pra/P ti/PRO <i>entrares/SR</i> numa/P+D-F faculdade/N
<b>Quiser</b>	... se/SE Deus/NPR <i>quiser/VB-SR</i> eu/PRO

**Tabela 45:** Verbo conjugado etiquetado como *SER* verbo infinitivo

Na tabela 46, estão as 7 ocorrências de verbos terminados em *-rei* etiquetados como verbos no futuro.

OCORRÊNCIA	CONTEXTO
<b>Adorei</b>	... porque/CONJ <i>adorei/VB-R</i> aquela/D-F terra/N ,/, o/D povo/N
<b>Colaborei</b>	... eu/PRO também/ADV <i>colaborei/VB-R</i> ./ Eu/PRO acho/VB-P
<b>Formei</b>	... eu/PRO me/CL <i>formei/VB-R</i> eu/PRO me/CL formei/VB-D em/P
<b>Morei</b>	... sempre/ADV <i>morei/VB-R</i> nessa/P+D-F aqui/ADV nessa/P+D-F casa
<b>Namorei</b>	... <i>namorei/VB-R</i> e/CONJ <i>noivei/VB-R</i> na/P+D-F Praça/NPR da/P+D-F
<b>Noivei</b>	... <i>namorei/VB-R</i> e/CONJ <i>noivei/VB-R</i> na/P+D-F Praça/NPR da/P+D-F
<b>Parei</b>	... quando/CONJS eu/PRO <i>parei/VB-R</i> de/P ir/VB pras/P+D-F-P

**Tabela 46:** Verbo terminado em *-rei* etiquetado como verbo futuro

Na tabela 47, estão as 17 ocorrências de verbos terminados em *-S* etiquetados como nomes no plural.

OCORRÊNCIA	CONTEXTO
<b>Botas</b>	... tu/PRO <i>pegas</i> /N-P o/D microfone/N e/CONJ <b>botas</b> /N-P no/P+D chão
<b>Chegas</b>	... tu/PRO <b>chegas</b> /N-P lá/ADV ./, já/ADV não/NEG dá/VB-P ./,
<b>Cozinhas</b>	... de/P noite/N tu/PRO <b>cozinhas</b> /N-P sempre/ADV ./, assim/ADV
<b>Desestruturamos</b>	... não/NEG <i>nos</i> /P+D-P <b>desestruturamos</b> /N-P ?/. Porque/CONJS a/D-F
<b>Encontras</b>	... ./, quando/ADV tu/PRO <b>encontras</b> /N-P <i>al</i> /ADV às/P+D-F vezes/N-P
<b>Entras</b>	... quando/CONJS tu/PRO <b>entras</b> /N-P no/P+D avião/N ./, aquele/D
<b>Escutas</b>	... tu/PRO não/NEG <b>escutas</b> /N-P <i>canis</i> /VB-P ./, têm/TR-P muito/Q
<b>Fechávamos</b>	... às/P+D-F-P duas/NUM horas/N <b>fechávamos</b> /N-P o/D programa/N
<b>Gastas</b>	... tu/PRO <b>gastas</b> /N-P quando/CONJS tu/PRO <i>trocas</i> /N-P de/P governo
<b>Levas</b>	... então/ADV tu/PRO <b>levas</b> /N-P muito/Q tempo/N
<b>Organizas</b>	... que/C tu/PRO <b>organizas</b> /N-P os/D-P lugares/N-P ./, o/D
<b>Passas</b>	... tu/PRO <b>passas</b> /N-P em/P <i>túnel</i> /NPR ./, ainda/ADV pode/VB-P
<b>Pedirias</b>	... mágica/ADJ ./, o/D que/WPRO que/C tu/PRO <b>pedirias</b> /N-P ?/. Três
<b>Pegas</b>	... tu/PRO <i>fazes</i> /VB-P ./, tu/PRO <b>pegas</b> /N-P o/D microfone/N
<b>Programas</b>	... tu/PRO te/CL <b>programas</b> /N-P ./, eu/PRO sei/VB-P que/C
<b>Queiras</b>	... Mas/CONJ tu/PRO não/NEG <b>queiras</b> /N-P saber/VB ./, rapaz/N
<b>Trocas</b>	... tu/PRO <i>gastas</i> /N-P quando/CONJS tu/PRO <b>trocas</b> /N-P de/P governo

Tabela 47: Verbo terminado em -S etiquetado como nome plural

Na tabela 48, está a única ocorrência de verbo etiquetado como preposição.

OCORRÊNCIA	CONTEXTO
<b>Arruma</b>	... aí/ADV eu/PRO <b>arruma</b> /P mala/N ./, e/CONJ pensar/VB e/CONJ

Tabela 48: Verbo etiquetado como preposição

Na tabela 49, está a única ocorrência de verbo presente etiquetado como verbo gerúndio.

OCORRÊNCIA	CONTEXTO
<b>Ajudo</b>	... com/P ela/PRO ./, <b>ajudo</b> /VB-G um/D-UM pouco/Q na/P+D-F casa/N

Tabela 49: Verbo presente etiquetado como verbo gerúndio

Na tabela 50, está a única ocorrência de verbo presente etiquetado como verbo imperativo.

OCORRÊNCIA	CONTEXTO
<b>Benze</b>	... qualquer/Q-G coisa/N ./, ele/PRO <i>benze</i> /VB-I ali/ADV ./, leva/VB-P

**Tabela 50:** Verbo presente etiquetado como verbo imperativo

Na tabela 51, está o único caso de verbo presente etiquetado como verbo infinitivo.

OCORRÊNCIA	CONTEXTO
<b>Animo</b>	... eu/PRO acho/VB-P que/C me/CL <i>animo</i> /VB só/FP fazendo/VB-G

**Tabela 51:** Verbo presente etiquetado como verbo infinitivo

Na tabela 52, está a única ocorrência de verbo indicativo etiquetado como verbo subjuntivo.

OCORRÊNCIA	CONTEXTO
<b>Engole</b>	... a/D-F gente/N <i>engole</i> /VB-SP muito/Q o/D esse/D

**Tabela 52:** Verbo indicativo etiquetado como verbo subjuntivo

Na tabela 53, estão as 3 ocorrências de verbos gerúndio etiquetados como verbos particípio.

OCORRÊNCIA	CONTEXTO
<b>Enrolando</b>	... pessoa/N assim/ADV <i>enrolando</i> /VB-AN feno/N ou/CONJ um/D-UM
<b>Incubando</b>	... ./, aí/ADV fica/VB-P <i>incubando</i> /VB-AN parece/VB-P que/C não
<b>Latindo</b>	... eu/PRO ouvi/VB-D cachorro/N <i>latindo</i> /VB-AN ./, aí/ADV eu/PRO

**Tabela 53:** Verbo gerúndio etiquetado como verbo particípio

Na tabela 54, está o único caso de verbo presente etiquetado como verbo particípio.

OCORRÊNCIA	CONTEXTO
<b>Frito</b>	... primeiro/ADV eu/PRO <i>frito</i> /VB-AN cebola/N ./, aí/ADV depois/ADV

**Tabela 54:** Verbo presente etiquetado como particípio

Na tabela 55, está a ocorrência de verbo passado etiquetado como verbo presente.

OCORRÊNCIA	CONTEXTO
Estudou	... quem/WPRO <i>estudou</i> /VB-P muito/Q ,/, quem/WPRO

**Tabela 55:** Verbo passado etiquetado como verbo presente

#### 4.3.1.2 Erros de anotação da fala

O Aelius, como mencionamos anteriormente, é um etiquetador treinado para anotação automática de *corpora* em língua escrita. Por isso, ele encontrou alguma dificuldade em analisar expressões típicas da fala, como marcadores conversacionais, hesitações e truncamentos. Abaixo, na tabela 56, destacamos os erros desse tipo que foram encontrados durante a correção manual, além de mostrar o contexto em que ocorreram. Na próxima seção, proporemos etiquetas que sejam adequadas para descrever essas expressões.

OCORRÊNCIA	CONTEXTO
<b>Ãh</b>	... como/CONJS é/SR-P que/C se/SE diz/VB-P ./, <b>ãh/N</b> Ah/INTJ !/ ... morar/VB num/P num/P centro/N mais/ADV-R <b>ãh/NPR</b> assim/ADV
<b>Ahã</b>	... isto/DEM <b>ahã/ADV</b> ./, são/SR-P daqui/P+ADV
<b>Bé</b>	... ovelhinha/N fazia/VB-D “/QT <b>bé/NPR</b> “/QT ./, “/QT <b>bé/NPR</b> “/QT
<b>Bom</b>	... eu/PRO digo/VB-P :/. <b>Bom/ADJ</b> ./, vou/VB-P ter/TR-P que/WPRO
<b>Claro</b>	... muito/Q gaúcho/N ./, <b>claro/ADJ</b> ./, <b>né/NPR</b> ?/.
<b>Hum</b>	... próximas/ADJ-F-P eleições/N-P ?/. <b>Hum/D-UM</b> ./, E/CONJ o/D
<b>Hum-hum</b>	... aqui/ADV do/P+D lado/N ?/. <b>Hum/D-UM hum/D-UM</b> ./, Inclusive
<b>Né</b>	... era/SR-D hoje/ADV o/D sem/P fio/N ./, <b>né/NPR</b> ?/.
<b>Olha</b>	... F/NPR <b>Olha/VB-P</b> ./, <b>chimarrão/VB-R</b> a/D-F gente/N fazia
<b>Tá</b>	... aí/ADV <b>tá/ADV</b> ./, mas/CONJ a/D-F guria/N nervosa/ADJ-F ... o/D cara/N ./, <b>tá/NPR</b> ./, tudo/Q o/D cara/N
<b>Tal</b>	... todo/Q mundo/N no/P+D banho/N e/CONJ <b>tal/ADJ-R-G</b>
<b>Tarará</b>	... trabalho/N ./, ((“/QT <b>tarará/NPR</b> “/QT ))/( ./, leva/VB-P trabalho/N
<b>Tipo</b>	... coisas/N-P lá/ADV assim/ADV <b>tipo/N</b> ./, quando/CONJS eu/PRO

Tabela 56: Erros de anotação da fala

Por se tratar de um *corpus* de língua falada, outro problema que o Aelius encontra ao etiquetar são os truncamentos, que é claro, geram etiquetas diversas, pois não há a etiqueta “truncamento”. Para mencionar alguns exemplos, temos pseudo-palavras como “caind”, “cans”, “di”, “el”, “mu”, “mour”, “maq”, “maqui”, “su”, e sub. Essas palavras receberam as mais variadas etiquetas: nome próprio, nome plural, verbo, advérbio, etc. Na tabela 57, mostraremos os contextos em que ocorreram:



TRUNCAMENTO	CONTEXTO
<b>caind</b>	... ./, acaba/VB-P caind/P <i>caindo</i> /VB-G em/P domínio/N público/ADJ
<b>cans</b>	... tínhamos/TR-D <i>cans</i> /N-P <i>calçamento</i> /NPR e/CONJ não/NEG
<b>di</b>	... tenha/VB-SR opções/N-P <i>di</i> /VB-P diferentes/ADJ-G-P ./, quero
<b>el</b>	... compensar/VB o/D fato/N de/P <i>el</i> /NPR dela/P+PRO não/NEG estar
<b>maq/maqui</b>	... <i>maquiar</i> /VB <i>maqui</i> /P+ADV maq/VB-D aí/ADV já/ADV é/SR-P
<b>mu/mour</b>	... nós/PRO fomos/VB-D <i>mu</i> /VB-D <i>mour</i> /VB fomos/VB-D morar
<b>su/sub</b>	... matérias/N-P foram/SR-D <i>su</i> /VB-D <i>sub</i> /N-P suprimidas/VB-NA-F-

Tabela 57: Exemplo de truncamento

Como podemos observar, o truncamento *cans* foi etiquetado como nome plural, possivelmente por terminar em *-s*. A palavra seguinte, *calçamento*, foi etiquetada como nome próprio, e isso provavelmente aconteceu porque as etiquetas se relacionam umas com as outras, ou seja, se antes da palavra *calçamento* houvesse um determinante e não um nome plural, a probabilidade de ser marcada corretamente como nome seria maior. Segundo nossa análise, aparentemente os outros exemplos de truncamento não influenciaram em etiquetas erradas em outros vocábulos (ver os demais exemplos na tabela 57).

Acreditamos que casos que lidam com truncamentos sejam um pouco mais difíceis de resolver, pois é bastante complicado que o etiquetador reconheça quando se trata de uma palavra ou de uma “pseudo-palavra”. Pensamos que uma maneira possível de resolver o problema com palavras truncadas possa ter alguma relação com o algoritmo elaborado por Alencar (2009). Nesse texto, o autor do Aelius propõe um reconhecedor de neologismos. O truncamento obviamente não é um neologismo, mas é uma “palavra” de certa maneira nova. O anotador deve ser capaz de reconhecer quando uma palavra aparece ao lado de um fragmento que repete sua estrutura, antecedendo-a (como no exemplo *su sua*) ou sucedendo-a (como no exemplo *maquiagem maqui maq*), que destacamos na tabela.

Importante mencionar que no *corpus* há algumas palavras com a ortografia incorreta, o que faz uma etiqueta errada ser gerada, por exemplo:

- ➔ *Automóvel*: está grafado como *automável*, gerando etiqueta de nome próprio (... *se/SE comprava*/VB-D *automável*/NPR *lá/ADV*).

→ *Consanguíneo*: está grafado como consangüíno, gerando etiqueta de nome próprio ( ... um/D-UM relacionamento/N consangüíno/NPR ./, mas/CONJ).

Além disso, há palavras que recebem a etiqueta adequada, e em outro contexto recebem uma etiqueta errada, para citar alguns exemplos:

→ *Brinquedo*: recebe etiqueta errada de adjetivo e de verbo participípio, e a correta, de nome

- *Brinquedo*: etiquetado como adjetivo

Exemplo) ... pensar/VB num/P+D-UM *brinquedo*/ADJ e/CONJ imaginar/VB um

- *Brinquedo*: etiquetado como verbo participípio

Exemplo) ... mais/ADV-R em/P fazer/VB *brinquedo*/VB-AN ./, Pelo/P+D que

- *Brinquedo*: etiquetado como **nome**

Exemplo) ... nem/CONJ-NEG pensam/VB-P em/P *brinquedo*/N ./, nem/CONJ-NEG

→ *Inglês*: recebe a etiqueta errada de adjetivo, e a correta, de nome

- *Inglês*: etiquetado como adjetivo

Exemplo) ... países/N-P falam/VB-P em/P *inglês*/ADJ ?/.

- *Inglês*: etiquetado como **nome**

Exemplo) ... noção/N muito/Q rápida/ADJ-F de/P *inglês*/N ./, de/P *francês*/ADJ

→ *Latim*: recebe a etiqueta errada de nome próprio, e a correta, de nome

- *Latim*: etiquetado como nome próprio

Exemplo) ... da/P+D-F língua/N dentro/ADV do/P+D *latim*/NPR ./, né/NPR ?/.

- *Latim*: etiquetado como **nome**

Exemplo) ...o/D *inglês*/ADJ o/D *francês*/ADJ e/CONJ o/D *latim*/N era/SR-D

obrigatórios/N-P

→ *Montevideo*: recebe a etiqueta errada de nome comum, e a certa, de nome próprio

- *Montevideo*: etiquetado como nome comum

Exemplo) ... seguidamente/ADV vou/VB-P a/P *Montevideo*/N ./, e/CONJ a/D-F

- *Montevideo*: etiquetado como **nome próprio**

Exemplo)... lá/ADV em/P *Montevideo*/NPR porque/CONJS tenho/VB-P parentes

→ *Persianas*: recebe a etiqueta errada de adjetivo, e a certa, de nome

- *Persianas*: etiquetado como adjetivo

Exemplo) ... ainda/ADV tem/TR-P aquelas/D-F-P *persianas*/ADJ-F-P né/NPR ?/.

- *Persianas*: etiquetado como **nome**

Exemplo)... agora/ADV é/SR-P tudo/Q com/P persiana/N ./, com/P além/ADV

→ **Princesa**: recebe a etiqueta errada de nome próprio, e a certa, de nome

- *Princesa*: etiquetado como nome próprio

Exemplo) ... E/CONJ a/D-F *princesa*/NPR boa/ADJ-F atendia/VB-D os/D-P pedidos

- *Princesa*: etiquetado como **nome**

Exemplo) ... E/CONJ aí/ADV a/D-F *princesa*/N ruim/ADJ-G dava/VB-D ./, batia

→ **Sacola**: recebe a etiqueta errada de advérbio, e a certa, de nome

- *Sacola*: etiquetado como advérbio

Exemplo) ... eu/PRO ia/VB-D com/P *sacola*/ADV fazer/VB compras/N-P de/P

- *Sacola*: etiquetado como **nome**

Exemplo) ... que/C a/D-F *sacola*/N a/D-F *sacola*/N vai/VB-P vazia/ADJ-F ./.

→ **Ser**: recebe a etiqueta errada de verbo ser infinitivo, e a certa, de nome

- *Ser*: etiquetado como verbo ser infinitivo

Exemplo) ... é/SR-P próprio/ADJ do/P+D *ser*/SR humano/ADJ ./.

- *Ser*: etiquetado como **nome**

Exemplo) ... serzinho/N se/SE transformando/VB-G em/D-UM *ser*/N humano/ADJ

→ **Tule**: recebe a etiqueta errada de verbo, e a certa, de nome

- *Tule*: etiquetado como verbo

Exemplo) ... ./, e/CONJ *tule*/VB-SP ./, né/NPR ?/. *Saia*/VB-SP de/P cetim/N com/P

- *Tule*: etiquetado como **nome**

Exemplo) ... *Saia*/VB-SP de/P cetim/N com/P *tule*/N por/P cima/ADV ./, né/NPR ?/.

→ **Vestibular**: recebe a etiqueta errada de verbo infinitivo, e a certa, de nome

- *Vestibular*: etiquetado como verbo infinitivo

Exemplo) ... quando/CONJS eu/PRO fiz/VB-D *vestibular*/VB pra/P faculdade/N

- *Vestibular*: etiquetado como nome

Exemplo) ... passei/VB-D *direto*/ADJ no/P+D *primeiro*/ADJ *vestibular*/N ./.

Esses casos de a mesma palavra receber etiquetas diferentes em contextos distintos, muito provavelmente, ocorrem por causa da relação com as etiquetas que estão ao redor da palavra, ou seja, que estão no mesmo *n*-grama.

No próximo capítulo, proporemos algumas sugestões para os casos analisados oriundos especificamente da anotação de língua falada, esquematizados na tabela 56.

## 5 SUGESTÕES PARA IMPLEMENTAÇÕES FUTURAS NO AELIUS

Nesta seção, vamos sugerir implementações para que o Aelius etiquete com ainda mais eficiência *corpus* de língua falada. Primeiramente, é importante advertir mais uma vez que se trata de um *corpus* de língua falada, pois, a acurácia do etiquetador depende totalmente do tipo do *corpus*. Para exemplificar, um etiquetador que foi construído a partir do treinamento de um *corpus* de textos da área da astronomia tem a tendência de diminuir o índice de acertos na etiquetagem de um *corpus* formado por textos sobre culinária<sup>1</sup>.

O Aelius – ou qualquer outro etiquetador –, não vai necessariamente acertar a etiqueta de uma mesma palavra em todas as suas ocorrências, pois qualquer mudança, por menor que seja, no contexto sintático da palavra a ser etiquetada pode bastar para que o etiquetador decida por outra etiqueta. Temos de levar em conta que a análise morfossintática que nós, linguistas, fazemos considera uma grande quantidade de regras e padrões que nos são naturais, mas, ao etiquetador, foram passadas apenas algumas dessas informações, através de moldes estatísticos; por isso, algumas vezes o etiquetador acaba sendo induzido a etiquetar equivocadamente, pois ele não conta com a intuição de falante natural.

### 5.1 Sugestões e novas etiquetas

Com base em toda a análise que fizemos da anotação do Aelius, percebemos que seria muito valiosa a inclusão de novas etiquetas, que pudessem fazer com que o índice de acertos do anotador fosse ainda mais alto. Para fazermos implementações em um etiquetador, devemos utilizar um *corpus* já anotado e revisado, para que se treine um algoritmo de aprendizagem de máquina. Os algoritmos utilizam-se de regras estatísticas para aprender como as etiquetas se relacionam entre si. Exemplificando, a frequência de determinantes antes de substantivos é muito grande, por isso, há grandes chances de a

---

<sup>1</sup> A esse respeito, cf. Alúísio e Almeida (2006)

palavra que vem antes de um nome ser um determinante, e vice-versa, e o anotador adquire essa noção através do que é “ensinado” a ele, com regras estatísticas, e com a inserção de palavras no dicionário do programa. Segundo Voutilainen (2009), *apud* Alencar (2012)

essas regras podem ser baseadas sobre duas fontes de informações, ambas codificadas no etiquetador na forma de uma linguagem modelo: a informação sobre a palavra em si, ou seja, em que contexto efetivo a palavra é mais usada, por exemplo, a palavra como: verbo ou advérbio; e as informações sobre a sequência da palavra palavra/etiqueta (ou contexto informacional): isto é, o modelo pode preferir analisá-lo como um verbo a uma conjunção, se o termo precedente for um advérbio ou um determinante.

Para a inserção de novas etiquetas, é necessário o treinamento de um algoritmo de aprendizagem de máquina com um *corpus* que tenha passado por anotação e revisão, com essas novas etiquetas, para que se arquitecte um novo etiquetador. Por esse motivo, aqui vamos sugerir novas etiquetas, além de sugerir que sejam inseridas novas palavras na lista de interjeições, baseando-nos em nossa análise linguística dos dados que levantamos a partir da correção manual sistematizados nas tabelas da seção anterior. Com a implementação dessas etiquetas, esperamos que a acurácia do Aelius seja ainda mais satisfatória em sua tarefa de anotação de língua falada.

### 5.1.1 Inserção de palavras na etiqueta interjeição

O etiquetador Aelius já possui em sua etiquetagem a classe das interjeições e etiqueta corretamente algumas, como por exemplo, “ah”, “hem” e “ó”. Porém, como vimos nas tabelas 15, 16, 17 e 18, há outras palavras da classe das interjeições que o Aelius não reconheceu corretamente e as etiquetou como adjetivo, nome, nome próprio e verbo. Para que isso seja resolvido, sugerimos que as palavras *Ai*, *Bah*, *Barbaridade*, *Chê*, *Credo*, *Hãn*, *Ô*, *Ôpa*, *Pô*, *Poxa*, *Puxa*, *Ué*, *Tchê* e *Xi* sejam acrescentadas ao dicionário do etiquetador, na classe das interjeições.

Como as palavras *Barbaridade* e *Credo* podem também ser nomes, e as palavras *Pronto* e *Puxa* podem ser verbos, é preciso que o algoritmo seja treinado para poder reconhecer as possibilidades de etiquetas da mesma palavra, com a devida atenção a esses casos. Acreditamos também que um ponto exclamativo após a interjeição pode auxiliar o anotador a etiquetar corretamente a palavra, pois a chance de ocorrer um ponto de exclamação após uma interjeição é grande.

### 5.1.2 Inserção de novas etiquetas

Para que o Aelius possa dar conta de todas as palavras do *corpus* analisado – e futuramente, de outros *corpora* de língua falada –, pensamos que é necessário que haja uma etiqueta que dê conta de casos muito específicos da fala, como as onomatopeias e aféreses, e além disso, o caso dos marcadores conversacionais, que discutiremos a seguir.

#### 5.1.2.1 Onomatopeias

Onomatopeia é a criação de uma palavra para reproduzir o mais fielmente possível um som. Em nossa análise, ocorreu apenas um caso de onomatopeia, a imitação do som de uma ovelha: *bé*. Para que o Aelius dê conta dessa e de outras onomatopeias que possam aparecer em *corpora* futuros, acreditamos que seja adequado inserir a etiqueta “**ONTP**” - Onomatopeia, para que o anotador etiquete corretamente as seguintes palavras: argh, atchim, au, bang, bé, blá, brr, bzz, clique, cof, grr, hum hum, nhac, tec, tic, tac, toc, plaft, piu, pof, pum e vrum. Com o passar do tempo, se surgir a necessidade, essa lista pode vir a aumentar. Esperamos que assim o Aelius dê conta desse recurso de formação de palavra, não muito produtivo, mas encontrado nas línguas naturais, especialmente na modalidade falada.

### 5.1.2.2 Aférese

A aférese é a supressão de fonema no início das palavras. No *corpus* do Varsul analisado, ocorreram apenas casos de aférese verbal, com os quais o etiquetador conseguiu lidar satisfatoriamente, como *tá*, *tão*, etc. No entanto, como pretendemos propor melhorias para que o Aelius possa analisar textos na modalidade falada da língua, de maneira irrestrita, acreditamos que seja válido preparar o anotador para lidar com casos de aférese nominal, adjetival, preposicional e adverbial, como por exemplo: “fessora” (professora), “brigada” (obrigada), té (até), “bora” (embora), etc.

### 5.1.2.3 Marcadores conversacionais

Os marcadores conversacionais são palavras ou expressões típicas e corriqueiras da fala, principalmente da conversa, que possuem um aumento em seu escopo de significados e funcionam como subsídios para a interação. Há marcadores utilizados pelo ouvinte e pelo falante, e os dois apoiam-se mutuamente para que o diálogo continue. Os marcadores são utilizados para marcar hesitação, servem como busca de apoio do interlocutor e funcionam como monitoramento do ouvinte. Segundo Urbano (2010: 93), os marcadores conversacionais são

elementos de variada natureza, estrutura, dimensão, complexidade semântico-sintática, aparentemente supérfluos ou até complicadores, mas de indiscutível significação e importância para qualquer análise de texto oral e para sua boa e cabal compreensão.

Esses marcadores podem ser linguísticos ou não. Os marcadores não linguísticos (ou “extralinguísticos”) são ações como risos e olhares, por exemplo, e não trataremos

deles aqui, pois não temos como captá-los, justamente por não serem marcas linguísticas *per se*<sup>2</sup>.

Os marcadores conversacionais linguísticos se dividem em prosódicos e verbais. Os marcadores prosódicos também não serão analisados neste trabalho, porque são os alongamentos de sílabas, as pausas e as mudanças na tessitura ou na velocidade da voz, ou seja, são aqueles aspectos que não estão marcados no *corpus* e só conseguiriam ser percebidos se ouvíssemos as entrevistas. Os marcadores verbais são aspectos que podemos notar na transcrição das entrevistas e se dividem em lexicalizados e não lexicalizados, segundo a classificação proposta por Urbano (2010). Tanto os lexicalizados como os não lexicalizados tiveram ocorrência em nosso *corpus*. Os lexicalizados são *Bom*, *Claro*, *Tipo*, etc., e os não lexicalizados são *Ãh*, *Ahã* e *Hum*, por exemplo. Urbano ainda fala sobre a diferença dos marcadores em relação a seu lugar na sentença, dizendo que alguns têm mais tendência a ocorrer no início, como *Bom* e *Olha*, outros têm mais tendência de ocorrer no meio da frase, como *E tal* e *Claro*, e outros, ainda, têm maior probabilidade de ocorrer ao final, como *Tá* e *Né*. Em nosso *corpus* tivemos ocorrências de todos esses tipos, e vamos analisá-las a seguir.

- Marcadores conversacionais não lexicais:

**Ãh**, **Ahã**, **Hum**, **Hum hum** e **Tarará** foram os casos que ocorreram em nosso *corpus*; por isso, sugerimos que eles sejam adicionados ao dicionário do etiquetador, com a etiqueta **MC**, de marcador conversacional. Além dessas ocorrências, sugerimos que sejam adicionadas palavras como **Eh**, **Ahn** e **Uhn**, que também podem aparecer em *corpora* futuros.

- Marcadores conversacionais lexicais:

**Bom**, **Claro**, **Né**, **Olha**, **Tá**, **Tal** e **Tipo** são ocorrências do nosso *corpus*; portanto, sugerimos que essas palavras sejam adicionadas com a etiqueta **MCL**,

---

<sup>2</sup> Na verdade, há registro de risos no *corpus* analisado, mas escrito como “*risos*”, e etiquetado como nome plural. Como as etiquetas se relacionam entre si, e levam em conta essa etiqueta de nome que foi marcada, sugerimos que nos próximos textos essa marca seja retirada do *corpus*.



de marcador conversacional lexical. Aqui há de se tomar mais cuidado, pois, como todas são palavras lexicais, têm mais possibilidades de significado; por isso, o algoritmo precisa ser bem treinado, levando em conta as que aparecem mais no início, as mediais, e as finais.

O estudo das posições dos marcadores é bastante importante, por exemplo, quando a ocorrência *Bom* acontece como primeira palavra da frase, seguida por uma vírgula ou ainda, se acontece no meio da sentença mas entre vírgulas, a chance de ser um adjetivo é muito pequena, e a chance de ser um marcador discursivo é muito grande.

Em comunicação pessoal, o professor Sergio Menezzi sugeriu que talvez os marcadores conversacionais que aparecem em posição final podem estar sendo confundidos com vocativo, por causa de sua posição na frase, ocorrendo sempre ao final e geralmente após uma vírgula, seguida por um ponto de exclamação ou interrogação, bem como ocorreria com um vocativo, que é usualmente um nome próprio, o que justificaria essa etiqueta. Exemplo: o trecho que encontramos *fazer rancho e tudo mais, né?* poderia facilmente ser substituído por um trecho com um nome próprio: “fazer rancho e tudo mais, João?”, ainda que a força ilocutiva da sentença tenha sido alterada – algo que passa despercebido pelo Aelius. Tanto isso parece ser assim que a análise feita pelo Aelius desse trecho foi a seguinte: *fazer/VB rancho/N e/CONJ tudo/Q mais/ADV-R ,/, né/NPR ?/*. Esse raciocínio deve ser levado em conta quando a posição dos marcadores for estudada.

Com essas sugestões, esperamos que a taxa de acerto do Aelius para etiquetar língua falada automaticamente, que já é alta, seja ainda mais satisfatória. No próximo capítulo, faremos algumas considerações finais sobre nosso trabalho.

## 6 CONSIDERAÇÕES FINAIS

A Linguística de Corpus vem evoluindo cada vez mais ao longo dos últimos anos, mas um problema frequentemente encontrado é que muitas vezes não há ferramentas grátis disponíveis para que os pesquisadores possam ter mais eficiência e versatilidade em seus objetivos de pesquisa. Com a finalidade de auxiliar pesquisadores em linguística o Aelius foi criado, e a fim de aumentarmos seu escopo de anotação automática morfossintática para além de língua escrita, abrangendo também língua falada, fizemos este trabalho, analisando exaustivamente e cuidadosamente o *corpus* de língua falada anotado automaticamente. Esperamos que nossa pesquisa possa contribuir de maneira satisfatória com a melhoria desse programa que pode auxiliar muitas pesquisas em nossa área. Além disso, esperamos que cada vez mais os profissionais da linguagem tenham conhecimento dessas ferramentas que estão a nossa disposição e que efetivamente as usem, pois são instrumentos valiosos para a pesquisa. A importância da Linguística de Corpus, em conjunto com a Linguística Computacional, pode ser constatada na “elaboração de teorias gramaticais formalmente mais consistentes e psicolinguisticamente mais realistas [...] e, assim, testar, com um grau de sofisticação que dificilmente poderia ser atingido por seres humanos, a adequação dos modelos postulados” (ALENCAR; OTHERO, 2011: 9).

Em nosso trabalho, iniciamos falando sobre o que é a Linguística de Corpus e o que é um *corpus*, dando exemplos de como um *corpus* pode ser analisado, focando principalmente no assunto de nossa pesquisa, a anotação automática morfossintática.

Depois, apresentamos a ferramenta utilizada, o Aelius, e o nosso *corpus*, constituído por trechos de transcrição de entrevistas do Banco de Dados do Varsul. Após apresentarmos isso, passamos para a parte mais exaustiva do trabalho, a análise dos dados obtidos com a correção manual da anotação automática feita pelo anotador Aelius, onde destacamos os principais problemas encontrados na correção e separamos as ocorrências por tipos de erros (confusão de etiquetas ou erros de anotação da fala), mostrando os contextos em que ocorreram.

Com a intenção de adequarmos o Aelius para outros *corpora* de língua falada, além de melhorar a anotação do *corpus* do Varsul, sugerimos algumas implementações

de etiquetas (onomatopeias, aférese e marcadores conversacionais) e adição de novas palavras ao dicionário do etiquetador, no grupo das etiquetas das interjeições.

Com isso esperamos que a ferramenta seja aperfeiçoada em uma próxima versão, e além disso, que seja mais utilizada por linguistas, pois esperamos que o Aelius obtenha maior número de acertos – não esquecendo que ele já tem uma acurácia alta – e que o Varsul tenha à sua disposição um etiquetador automático morfossintático de qualidade, para futuras pesquisas.

Acreditamos que essas ferramentas computacionais que auxiliam o estudo da língua poderiam ser ainda mais usadas por linguistas porque há muitos *corpora* disponíveis e as ferramentas estão cada vez mais acessíveis, tornando nossas pesquisas mais eficientes e confiáveis.

## 7 REFÊRENCIAS

ALENCAR, Leonel Figueiredo de. **Aelius: uma ferramenta para anotação automática de corpora usando NLTK**. IX Encontro de Linguística de Corpus. Porto Alegre, PUCRS, 8 e 9 de outubro de 2010.

ALENCAR, Leonel Figueiredo de. **Superando o estado da arte na etiquetagem morfossintática por meio de regras de pós-etiquetagem**. *In: Anais do X Encontro de Linguística de Corpus – Aspectos metodológicos dos estudos de corpora*. Belo Horizonte: UFMG, 2012.

ALENCAR, Leonel Figueiredo de. Novos recursos do Aelius para o processamento computacional raso do português. *In: LAPORTE, E.; SMARSARO, A.; VALE, O. (Orgs.). Dialogar é preciso: linguística para o processamento de línguas*. Vitória: PPGEL/UFES, 2013.

ALENCAR, Leonel Figueiredo de; OTHERO, Gabriel de Ávila (Orgs.). **Abordagens computacionais da teoria da gramática**. Campinas: Mercado de Letras, 2011.

ALUÍSIO, Sandra Maria; ALMEIDA, Gladis M. B. O que é e como se constrói um corpus? Lições aprendidas na compilação de vários corpora para pesquisa linguística. **Calidoscópico**, vol. 4, n. 3, p. 156-178, set/dez 2006.

BECHARA, Evanildo. **Moderna Gramática Portuguesa**. Rio de Janeiro: Lucerna, 37ª edição, 2009.

BERBER SARDINHA, Tony. Linguística de corpus: histórico e problemática. *Revista D.E.L.T.A.*, vol. 16, N. 2, 2000a.

BERBER SARDINHA, Tony. O que é um corpus representativo? **DIRECT Papers** 44. São Paulo e Liverpool: LAEL & AELSU, 2000b.

BERBER SARDINHA, Tony. **Linguística de Corpus**. Barueri, SP: Manole, 2004.

BERBER SARDINHA, Tony. Linguística de Corpus: Uma entrevista com Tony Berber Sardinha, *Revista Virtual de Estudos da Linguagem - ReVEL*. Vol. 2, n. 3, agosto de 2004.

BRANCO, António e João Silva, 2004. **Evaluating Solutions for the Rapid Development of State-of-the-Art POS Taggers for Portuguese**. *In* Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa and Raquel Silva (orgs.), *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC2004)*, Paris, ELRA, ISBN 2-9517408-1-6, pp.507-510.

CASTILHO, Ataliba T. de. **A língua falada no ensino de português**. São Paulo: Editora Contexto, 2001.

COELHO, Izete Lehmkuhl. A ordem V DP em construções monoarumentais: uma restrição sintático-semântica. **Letras de Hoje**, 119, 2000.

DEZOTTI, Lucas Consolin. As “partes da oração” de Donato aos modistas. **ReVEL**, vol. 8, n. 14, 2010.

GARSDIE, Roger; LEECH, Geoffrey; McENERY, Anthony. **Corpus annotation: linguistic information from computer text corpora**. London / New York: Longman, 1997.

MITTMANN, Maryualê M. **Construções de alçamento a sujeito: uma análise com base em corpus de fala e escrita**. Dissertação de Mestrado. Florianópolis: UFSC, 2006.

OTHERO, Gabriel de Ávila; GASPERIN, Caroline Varaschin. **Linguista x computador - trabalhando com corpus eletrônico**, (manuscrito), 2001.

OTHERO, Gabriel de Ávila; MENUZZI, Sérgio de Moura. **Linguística Computacional teoria & prática**. São Paulo: Parábola, 2005.

RASO, Tommaso; MELLO, Heliana (Orgs.). **C-ORAL BRASIL I Corpus de referência de português brasileiro falado informal**. Belo Horizonte: Editora UFMG, 2012.

SCHWINDT, Luiz Carlos da Silva. **A harmonia vocálica em dialetos do sul do país: uma análise variacionista**. Dissertação de Mestrado. Porto Alegre: PUCRS, 1995.

TAVARES, Maria Alice. **Um estudo variacionista de aí, daí, então e e como conectores seqüenciadores retroativo-propulsores na fala de Florianópolis**. Dissertação de Mestrado. Florianópolis: UFSC, 1999.

URBANO, Hudinilson. **Marcadores Conversacionais**. In: Análise de textos orais. São Paulo: Humanitás, 2010.

## 8 ANEXOS

### 8.1 Texto cru

\*Olha, fiz Rio de Janeiro, Canecão, me apresentei no Canecão, no Hilton Hotel, fiz São Paulo, fiz Minas Gerais, fiz Juiz de Fora, né? \*Eu viajei bastante mesmo, viajei tranqüilo, aqui no Rio Grande do Sul, viajei quase todo o estado do Rio Grande do Sul, <viaj->. \*Era eu, o Carlitos Magallanes, o Rubens Val, nós tínhamos um grupo de típica mesmo, [típica]- (hes) típica clássica, minha dança era típica clássica. \*Eles se apresentavam como músicos e eu então fazia o bailarino, e ali [eu]- era uma vida boa. \*E quem dançava com o senhor? \*Era uma menina. \*Sempre ela? F\*[Sempre foi a <mesm->]- [não]- não, teve diversas. \*Tive diversas (inint) <part0ners>, |como se chama#se|, né? \*[Na]- (est) na parte artística, |o nome a ela se dá <part0ner>| eu # <o-> [a]- ensinei muitas meninas. \*Agora já não dá mais que estou muito barrigudo. \*(risos geral) [ (inint). ] assim alguma coisa engraçada ou 1 E alguma coisa interessante? \*Ah! \*Teve, teve uma história, até casualmente um rapaz muito meu amigo faleceu. \*[Nós fomos]- [nós estávamos]- eu ia fazer um show no Teatro Presidente, e casualmente veio o Pablo Sebastian0 a Porto Alegre, e num show dele, ele precisava de um casal de bailarinos, e o empresário dele0 me conhecia, casualmente o meu escritório era no mesmo prédio0 no que ele também tinha um escritório, aí, convidou: + "\*Olha, Cláudio, quem sabe tu aceitas, né? sair com ele." "\*É, né? não quero mais sair e 0 coisa." "\*Não mas quem sabe e coisa vamos lá e <coi->." "\*Tudo bem, quem é o cara?" "\*É o Pablo Sebastian." "\*Está tudo bem não tem problema." \*Aí fomos. \*Fomos pro Teatro Presidente. \*Aí chegamos lá no teatro0 à tardezinha, eram umas três e meia da tarde, para olharmos, observarmos, né? o palco pra nós termos uma distância entre músico, né? cantor0 e a minha pessoa e a minha <part0ner>, então aí, olha, quando0 chamei o Pablo Sebastian, disse pra ele: "\*Olha, Pablo, quando tu pegares o microfone- "\*Porque não era o microfone, [era sem]- era com fio, não era hoje o sem fio, né? "\*Tu não te esqueças de dar um0 laço0 para [sair]- tirar ele aqui do palco, porque as lâmpadas ficam escuras, [não]- [os]- os canhões, né?" (hes) batem no meu rosto, e eu não enxergo o chão, né? \*Então, a casa estava lotada, sabe? \*E esse meu amigo [estava]- era o primeiro da platéia pra bater palmas, (ruído) né? \*Ele gostava muito de ir, então, [todas]- todos os lugares que eu ia com ele, ele ia junto comigo, então eu usava um chapéu, um lenço, né? um chapéu, o típico mesmo, né? típico, então eu dava o chapéu a ele, então ele ia pro camarim0 quando eu trocava de chapéu, aquela coisa toda, então ele

sempre ia comigo, então ele era o primeiro, nesse dia ele foi e em vez de ir para o camarim, não, ele veio e sentou na platéia, ele queria0 assistir, né? ainda mais o Pablo Sebastian, conhecido, um nome famoso, né? aí tá0 e nesse meio tempo o Pablo Sebastian se apresentou, e se esqueceu0 de puxar o fio, não? \*E eu venho dançando pah! e coisa e o fio era grosso, o sapato escorrega e eu vou pular na platéia, meu filho, (inint) mas eu não sei como foi o espírito0 que eu ali eu consegui, [quando eu]- me desviar e 0levantar0 a guria e dançar com ela no ar, e este rapaz se 0levantou (ruído) ("nunca vi!") (risos geral) quer dizer, foi um caso de (inint). \*Outro foi na Tânia Carvalho, [nós <trabalhava->]- a Tânia Carvalho fazia programa no canal dez, então todas as quartas#feiras0 nós fazíamos [um]- uma entradinha, que era [a]- - a quinze pra- \*(hes) Uma e quarenta e cinco : às duas horas, ("fechávamos") o programa dela, então todas as quartas#feiras nós éramos ("tinha i era o contrato já firmado") \*(inint) era eu, o Rubens Val, Célia Franco, (toss) e um dia, eu disse pro rapaz0 (inint) que arrumava o cenário, disse: "\*Olha, [eu]- eu estou com um sapato novo0 (hes) e tu fazes o seguinte, tu pegas um microfone0 e botas no chão, que quando eu vou patinar, o pessoal pode0 dar a audição, né? dos meus pés". \*E o cara: "\*Tudo bem", né? \*Arrumou ali e tal e coisa, né? \*Deixou tudo prontinho0 (inint) vou dançar e tal e coisa, né? \*Rodando, né? \*Programa ao vivo, né? \*A guria de lá (inint). \*Eu entrava, depois entrava ela, eu me lembro que eu tinha que dar uma bofetada nela, era machão, ela tinha que apanhar, pra platéia gostar, né? eu vou fazer e ela não se dá conta que eu tinha dito pra ele botar o microfone, e ela vai e diz: "\*Olha o microfone!" \*(falando rindo) Esses - são os casos, entendeu? que acontecem (f), né? \*Outra vez foi no Rio de Janeiro, (inint) Hilton Hotel0 e0 todo o artista, quase oitenta por cento, vou dizer oitenta, me deu um estado # de nervos, ("que tu sabes") uma platéia é uma platéia, né? \*Então o cara tomou um traguinho ("vamos pegar") uma garrafinha de uísque, essas de bolso, né? botava escondidinha e tal, né? \*E chegava no camarim, se enfeitava, botava talco tal e coisa, por causa dos refletores e aquela coisa toda, então dava uma bicada, e eu fui e dei uma bicada, mas eu estava louco pra urinar, mas estava louco pra urinar mesmo, e não havia meio, rapaz, de saber onde era o tal de0 |banheiro pra mim ir|, né? onde era o WC que eu não sabia nada, né? \*("Enxerguei") um cara lá: "\*Vem cá, onde é-" \*O cara diz: "\*Não, eu também não sei, estou chegando agora, junto contigo." \*Né? \*E eu, tá, aí peguei a garrafa (inint) mizei, <puf-> fechei e deixei ali, né? \*A guria entra0 do outro camarim, vê a garrafa, pega e <puf!> ( risos geral). \*É fato que acontece. \*Outra vez foi [aqui em <Sã->]- [aqui em]- (ruído) em Jaguarão, nunca tinha ido a Jaguarão, não

conhecia, aí tá, vamos embora pra Jaguarão, então vamos, ("aí tá"). \*[O]- o ("Bisquintero"), que era o empresário, e o Salim (toss): "\*Olha, vocês vão fazer um show lá, tá? \*Termina o show, |pegas a kombi, se manda que vocês| vão ter que estar em Bagé hem! não esquece." "\*Não, tudo bem. \*Ma0s é uma apresentação só?" "\*Não, é uma apresentação só." "\*Tudo bem, vamos embora." \*Aí chegou de manhã, <pá> daqui, <pá> dali, <pá> <dacá>, chegou0 o ("Rudi") e disse pra mim: "\*Olha, Cláudio, a - turma0 que ia0 pifou, ah! \*Ma0s o show vai sair, eu entrei em contato0 com Montevideo, e vão ir os músicos pra lá, vão esperar vocês lá." "\*Então tudo bem." \*Mas eu não conhecia nenhum deles, né? \*Nem sabia quem era Tupamaro, quem deixava de ser Tupamaro, não sabia quem era ninguém. \*Aí pegamos a guria, pegamos esse rapaz, que eu falei, que se Olevantou, né? \*O falecido Dudu, o apelido dele0 era Dudu. "\*Dudu, vamos embora." "\*Então vamos embora." \*Aí fomos embora. \*Chegamos lá, uma terra que Deus o livre, bah! \*Aí tá, descemos na rodoviária, rodamos e rodamos, aí vamos procurar um hotel, aí o rapaz0 que estava [na <ro->]- na rodoviária nos esperando, chegou pra nós: "\*Olha, tem uma zebra pra vocês0s." \*Digo: "\*Aí, pára aí, outra. \*Qual é?" "\*Os hotéis |estão tudo lotados|. \*Né? \*E eu digo: "\*Ma0s e agora, meu filho?" "\*Não te preocupas, vamos dar um jeito0 (ruído) ("nós vamos arrumar umas dessas pensões aí"), dormitório0 dos caminhoneiros e0 que é pra quebrar o galho só, né? \*Pra descansar aí um pouco e tal e coisa, né? depois nós vamos seguir viagem." "Ah! \*Tudo bem. \*Não tem problema, vamos embora." \*Aí tá, ma0s a guria nervosa que estava, sabe? \*A guria novinha, nervosa, aí: "Tá, vamos pra lá." \*Aí fomos, chegamos lá na pensão lá. \*A senhora prontamente arrumou um quarto de casal0 e um quarto de solteiro pra ele. \*Nem perguntou se éramos casados ou se não éramos, né? viu é casado. \*Então tá, tudo bem, fomos pra lá, aí tá ("certo") ficamos ali, |se arrumamos|, trocamos de roupa, aquela coisa toda, tomamos banho. "\*Olha, vamos jantar, né? \*|Vamos se preparar| pra jantar e coisa e vamos até o clube, pra conhecer o clube." \*O show era à meia#noite, né? \*Primeiro entrava o conjunto pra depois entrar a típica, a típica sempre [é]- é a última, né? \*Aí tá, aí fomos pro restaurante, mas0 um restaurante0 lindo, dois pisos. \*Chegamos no restaurante e era uma gurizada, ma0s gurizada, mas era gurizada mesmo, sabe o que que era? \*Tudo : pivetezinho, mas gurizada mesmo. \*|Isso eu nunca me esqueço| na minha vida, nunca vou esquecer também, ma0s tudo gurizada, era tomando Coca#Cola, tomando cerveja ("essa coisa") e nós entramos, |fomos lá pra cima, jantarmos|. \*Aí tá, estamos na mesa ali (inint) olha pra mim (inint), ele olha pra mim, o meu companheiro esse, e diz: "\*Bah! Cláudio, ah! vamos embora sabe que |tu já



pensou|, (est) |tu já ouviu| quando |tu entrou| |esses guris falando tudo que| já estão com o convite na mão pra ir pro baile, e você0s dançar tango e esses guris aí só têm agora o iê#iê#iê deles. \*Bah! \*Vai ser horrível." \*Aí, a guria0 tremia. "\*Vamos firme, né?" \*Aí, tá, sentamos numa mesa e coisa, pedimos a janta, aí nesse meio tempo nós ficamos jantando, aí chegou [um]- um rapaz que é, casualmente é gerente0 : da Caixa Econômica Federal e se dava comigo aqui no Centro, foi companheiro meu aqui do Centro também, aí quando- "\*Olha, Cláudio, tal e coisa e <bá#bá#bá>, o que está fazendo aqui?" \*Aí contei pra ele, né? \*Digo: + 1 "\*Bah! \*Eu estou fazendo isso." "\*O quê? \*Não pode ser!" "\*É." \*("Disse"): "\*Bah! \*Mas não é possível, é tu o bailarino, bah! \*Mas tu não queiras saber, rapaz, bah! \*Isso aqui já está tudo espalhado0 tá, mas espalhou mesmo, diz que vem um casal de bailarinos fora de série, propaganda aí, bah! \*Mas tem assim que vou te contar, né? tu vais no clube daqui ah! tu vais ver0 que tem de propaganda!" \*Digo: "\*Bah!" \*E a guria tremia como uma capivara, aí tudo bom, aí jantamos (inint) tinha um guia nosso, né? que era empregado (toss) [do <em->]- do empresário, aí fomos pra o clube, bah! \*E chegamos no clube (inint) parede e tudo e cheia de cartaz e <bá#bá#bá>, aí a guria mais tremia ainda, (toss) aí entramos pra uma sala lá da diretoria, não tinha camarim, não tinha nada (inint) não tem mesmo, aí entramos, estavam os caras ali, <a-> arrumando0 os aparelhos deles, né? "\*Oi, oi, oi. \*Ah! + \*Sou brasileiro." "\*Ah! \*É brasileiro?" "\*Sou brasileiro." \*("Usted") é bailarino de tango? (inint) por favor." "\*Sim (inint) porquê?" \*Bom, aí ele ficou meio assim, né? \*O castelhano e ela: "Ai, ai, ai Cláudio, vamos embora, |vamos se mandar|." \*Então [ela tinha] ela mudava de roupa, dois trajés, né? \*(inint) depois outra, aí, fomos na mala0 e cadê a roupa? \*Faltava a saia, a blusa tinha, ma0s a saia não tinha, a saia era rasgada até aqui, quer dizer, não era rasgada, ("era costurada, né?") \*Era rasgada até aqui. \*Cadê a saia? \*Não tinha. \*E agora (inint) tudo, né? \*Conforme for vamos. \*Aí os caras estão lá, com os instrumentos, tocando e coisa, né? \*(inint) afinando0 os instrumentos deles, aí o cara larga uma música lá e eu pá! \*Gritei o nome da música (ele dá-lhe) e eu pá! \*Daqui a pouco se 0levantou um Tupamaro, ma0s era um baita Tupamaro: "\*<O! Muchacho yo soy el cantor, usted no é nada cá">, (inint)." \*Aí, nós tivemos que sair correndo, né? \*Aí 0 saímos. \*A guria (ruído) disparou. \*Fui |encontrar ela| na pensão, que a pensão ficava0 umas quatro quadras ("numa") rua reta lá. \*Era a única [que <ta->]- que era calçada, também, né? \*Aí fui lá: "\*Vamos, que é isso?" "\*Não, mas está louco, tu vais brigar com o homem." "\*Não, que nada, vamos com calma." "\*Não, mas vai dar zebra." "\*Não vai dar zebra, vamos embora." \*Aí tá, aí |se

arrumamos| tal e coisa, fomos pra lá, levamos as malas, as roupas, né? \*Que tinha que levar, [aí]- aí cheguei pro cara e disse: "\*Olha, o negócio é o seguinte": (ruído) tinha0 dois que tocavam bandonion, (toss) disse: "\*Olha, o problema é o seguinte, quando eu fizer com o dedo pra cima (toss)0 tu0 rasgas o bandonion, quer dizer, eu estou ouvindo a música0 e sei o exato momento, né? que ele vai fazer0 o babado. \*Aí tá, tudo bem, aí tá tal e coisa fui lá, entramos no salão. \*Bah! \*(ruído) Todo mundo, e aplaudiu, né? \*Nessas alturas estava tudo bem, e as luzes ficaram |meia0 Ofoscas|, e nós saímos dançando, um baita de um salão, todo mundo sentadinho bonito, sabe? \*Coisa mais linda, todo mundo sentado, aí tá saímos dançando e coisa, daqui a pouco eu vou, dou uma volta nela e <pim!> um tapa na bunda dela, desculpe o termo, né? \*Mas <pim!> tapa na bunda, daí quando eu dou o tapa- "\*Dá-lhe machão!" \*Aí o pessoal, entendeu? ficou mais à vontade, aí tá, deu tal e coisa e eu fiz sinal pro cara, o cara rasga e eu, deito ela, né? na perna, né? \*Ah! \*E o pessoal sai aplaudindo. \*Terminou e- "\*Vamos sair pro corredor", digo: "vamos sair." \*Mas o pessoal de cima0 |de invés de| ficar sentadinho, e nós não estávamos prestando atenção, vieram todos pra copa, ma0s toda a gurizada, aí <no-> aí a gurizada queria meu chapéu, queria a minha roupa, queria [um pedaço de um]- um pedaço de outra coisa, mas nós saímos apavorados, escoltados até.

### 8.1 Texto anotado automaticamente pelo Aelius

Olha/VB-I ./, fiz/VB-D Rio/NPR de/P Janeiro/NPR ./, Canecão/NPR ./, me/CL apresentei/VB-D no/P+D Canecão/NPR ./, no/P+D Hilton/NPR Hotel/NPR ./, fiz/VB-D São/NPR Paulo/NPR ./, fiz/VB-D Minas/NPR-P Gerais/NPR-P ./, fiz/VB-D Juiz/NPR de/P Fora/NPR ./, né/NPR ?/.

Eu/PRO viajei/VB-D bastante/ADJ-R-G mesmo/ADJ ./, viajei/VB-D tranqüilo/NPR ./, aqui/ADV no/P+D Rio/NPR Grande/ADJ-G do/P+D Sul/NPR ./, viajei/VB-D quase/ADV todo/Q o/D estado/N do/P+D Rio/NPR Grande/ADJ-G do/P+D Sul/NPR ./, viaj/NUM ./.

Era/SR-D eu/PRO ./, o/D Carlitos/NPR Magallanes/NPR ./, o/D Rubens/NPR Val/NPR ./, nós/PRO tínhamos/TR-D um/D-UM grupo/N de/P típica/NPR mesmo/ADJ ./, típica/ADJ-F típica/N clássica/ADJ-F ./, minha/PRO\$-F dança/N era/SR-D típica/N clássica/ADJ-F ./.

Eles/PRO se/SE apresentavam/VB-D como/CONJS músicos/N-P e/CONJ eu/PRO então/ADV fazia/VB-D o/D bailarino/N ./, e/CONJ ali/ADV eu/PRO era/SR-D uma/D-UM-F vida/N boa/ADJ-F ./.

E/CONJ quem/WPRO dançava/VB-D com/P o/D senhor/NPR ?/.

Era/SR-D uma/D-UM-F menina/N ./.

Sempre/ADV ela/PRO ?/.

Fsempre/ADJ-G foi/SR-D a/D-F mesm/N não/NEG não/N ./, teve/TR-D diversas/ADJ-F-P ./.

Tive/TR-D diversas/ADJ-F-P partners/N-P ./, como/CONJS se/SE chamase/VB-SD ./, né/NPR ?/.

Na/P+D-F na/P+D-F parte/N artística/ADJ-F ./, o/D nome/N a/P ela/PRO se/SE dá/VB-P partner/VB eu/PRO o/D a/CL ensinei/VB-D muitas/Q-F-P meninas/N-P ./.

Agora/ADV já/ADV não/NEG dá/VB-P mais/ADV-R que/C estou/ET-P muito/Q barrigudo/VB-AN ./.

./.

Assim/ADV alguma/Q-F coisa/N engraçada/VB-AN-F ou/CONJ E/CONJ alguma/Q-F coisa/N interessante/ADJ-G ?/.

Ah/INTJ !/.

Teve/TR-D ./, teve/TR-D uma/D-UM-F história/N ./, até/P casualmente/ADV um/D-UM rapaz/N muito/Q meu/PRO\$ amigo/N faleceu/VB-D ./.

Nós/PRO fomos/SR-D nós/PRO estávamos/ET-D eu/PRO ia/VB-D fazer/VB um/D-UM show/N no/P+D Teatro/NPR Presidente/NPR ./, e/CONJ casualmente/ADV veio/VB-D o/D Pablo/NPR Sebastian/NPR a/P Porto/NPR Alegre/NPR ./, e/CONJ num/P+D-UM show/N dele/P+PRO ./, ele/PRO precisava/VB-D de/P um/D-UM casal/N de/P bailarinos/N-P ./, e/CONJ o/D empresário/N dele/P+PRO me/CL conhecia/VB-D ./, casualmente/ADV o/D meu/PRO\$ escritório/N era/SR-D no/P+D mesmo/ADJ prédio/N no/P+D que/WPRO ele/PRO também/ADV tinha/TR-D um/D-UM escritório/N ./, aí/ADV ./, convidou/VB-D :/. "/QT Olha/VB-P ./, Cláudio/NPR ./, quem/WPRO sabe/VB-P tu/PRO aceitas/VB-AN-F-P ./, né/NPR ?/.

Sair/VB com/P ele/PRO ./.

"/QT "/QT É/SR-P ./, né/NPR ?/.

Não/NEG quero/VB-P mais/ADV-R sair/VB e/CONJ coisa/N ./.

"/QT "/QT Não/NEG mas/CONJ quem/WPRO sabe/VB-P e/CONJ coisa/N vamos/VB-P lá/ADV e/CONJ coi/N ./.

"/QT "/QT Tudo/Q bem/ADV ./, quem/WPRO é/SR-P o/D cara/N ?/.

"/QT "/QT É/SR-P o/D Pablo/NPR Sebastian/NPR ./.

"/QT "/QT Está/ET-P tudo/Q bem/ADV não/NEG tem/TR-P problema/N ./.

"/QT Aí/ADV fomos/VB-D ./.

Fomos/SR-D pro/P+D Teatro/NPR Presidente/NPR ./.

Aí/ADV chegamos/VB-P lá/ADV no/P+D teatro/N à/P+D-F tardezinha/N ./, eram/SR-D umas/D-UM-F-P três/NUM e/CONJ meia/ADJ-F da/P+D-F tarde/N ./, para/P olharmos/VB-SR ./, observarmos/VB-SR ./, né/NPR ?/.

O/D palco/ADJ pra/P nós/PRO termos/TR-F uma/D-UM-F distância/N entre/P músico/N ./, né/NPR ?/.

Cantor/N e/CONJ a/D-F minha/PRO\$-F pessoa/N e/CONJ a/D-F minha/PRO\$-F partner/NPR ./, então/ADV aí/ADV ./, olha/VB-P ./, quando/CONJS chamei/VB-D o/D Pablo/NPR Sebastian/NPR ./, disse/VB-D pra/P ele/PRO :/. "/QT Olha/VB-P ./, Pablo/NPR ./, quando/CONJS tu/PRO pegares/VB-SR o/D microfone/N "/QT Porque/CONJ não/NEG era/SR-D o/D microfone/N ./, era/SR-D sem/P era/N com/P fio/N ./, não/NEG era/SR-D hoje/ADV o/D sem/P fio/N ./, né/NPR ?/.

"/QT Tu/PRO não/NEG te/CL esqueças/VB-SP de/P dar/VB um/D-UM laço/N para/P sair/VB tirar/VB ele/PRO aqui/ADV do/P+D palco/N ./, porque/CONJ as/D-F-P lâmpadas/N-P ficam/VB-P escuras/ADJ-F-P ./, não/NEG os/CL os/D-P canhões/N-P ./, né/NPR ?/.

"/QT Batem/VB-P no/P+D meu/PRO\$ rosto/N ./, e/CONJ eu/PRO não/NEG enxergo/VB-P o/D chão/N ./, né/NPR ?/.

Então/ADV ./, a/D-F casa/N estava/ET-D lotada/VB-AN-F ./, sabe/VB-P ?/.

E/CONJ esse/D meu/PRO\$ amigo/N estava/ET-D era/SR-D o/D primeiro/ADJ da/P+D-F platéia/N pra/P bater/VB palmas/N-P ./, né/NPR ?/.

Ele/PRO gostava/VB-D muito/Q de/P ir/VB ./, então/ADV ./, todas/Q-F-P todos/Q-P os/D-P lugares/N-P que/WPRO eu/PRO ia/VB-D com/P ele/PRO ./, ele/PRO ia/VB-D junto/ADV comigo/P+PRO ./, então/ADV eu/PRO usava/VB-D um/D-UM chapéu/N ./, um/D-UM lenço/N ./, né/NPR ?/.

Um/D-UM chapéu/N ./, o/D típico/ADJ mesmo/ADJ ./, né/NPR ?/.

Típico/NPR ./, então/ADV eu/PRO dava/VB-D o/D chapéu/N a/P ele/PRO ./, então/ADV ele/PRO ia/VB-D pro/P+D camarim/N quando/CONJS eu/PRO trocava/VB-D de/P chapéu/N ./, aquela/D-F coisa/N toda/Q-F ./, então/ADV ele/PRO sempre/ADV ia/VB-D comigo/P+PRO ./, então/ADV ele/PRO era/SR-D o/D primeiro/ADJ ./, nesse/P+D dia/N ele/PRO foi/VB-D e/CONJ em/P vez/N de/P ir/VB para/P o/D camarim/N ./, não/NEG ./, ele/PRO veio/VB-D e/CONJ sentou/VB-D na/P+D-F platéia/N ./, ele/PRO queria/VB-D assistir/VB ./, né/NPR ?/.

Ainda/ADV mais/ADV-R o/D Pablo/NPR Sebastian/NPR ./, conhecido/VB-AN ./, um/D-UM nome/N famoso/ADJ ./, né/NPR ?/.

Aí/ADV tá/ADV e/CONJ nesse/P+D meio/ADJ tempo/N o/D Pablo/NPR Sebastian/NPR se/SE apresentou/VB-D ./, e/CONJ se/SE esqueceu/VB-D de/P puxar/VB o/D fio/N ./, não/NEG ?/.

E/CONJ eu/PRO venho/VB-P dançando/VB-G pah/INTJ !/.

E/CONJ coisa/N e/CONJ o/D fio/N era/SR-D grosso/ADJ ./, o/D sapato/N escorrega/VB-P e/CONJ eu/PRO vou/VB-P pular/VB na/P+D-F platéia/N ./, meu/PRO\$ filho/N ./, mas/CONJ eu/PRO não/NEG sei/VB-P como/WADV foi/SR-D o/D espírito/N que/WPRO eu/PRO ali/ADV eu/PRO consegui/VB-D ./, quando/CONJS eu/PRO me/CL desviar/VB e/CONJ levantar/VB a/D-F guria/N e/CONJ dançar/VB com/P ela/PRO no/P+D ar/N ./, e/CONJ este/D rapaz/N se/SE levantou/VB-D (( "/QT nunca/ADV-NEG vi/VB-D !/.

"/QT ))( Quer/VB-P dizer/VB ./, foi/SR-D um/D-UM caso/N de/P ./.

Outro/OUTRO foi/SR-D na/P+D-F Tânia/NPR Carvalho/NPR ./, nós/PRO trabalhava/VB-D a/P Tânia/NPR Carvalho/NPR fazia/VB-D programa/N no/P+D canal/N dez/NUM ./, então/ADV todas/Q-F-P as/D-F-P quartasfeiras/ADJ-F-P nós/PRO

fazíamos/VB-D um/D-UM uma/D-UM-F entradinha/N ./, que/WPRO era/SR-D a/D-F a/P quinze/NUM pra/P Uma/D-UM-F e/CONJ quarenta/NUM e/CONJ cinco/NUM :/. às/P+D-F-P duas/NUM-F horas/N-P ./, (( "/QT fechávamos/N-P "/QT ))( o/D programa/N dela/P+PRO ./, então/ADV todas/Q-F-P as/D-F-P quartasfeiras/N-P nós/PRO éramos/SR-D (( "/QT tinha/TR-D i/N era/SR-D o/D contrato/N já/ADV firmado/VB-AN "/QT ))( era/SR-D eu/PRO ./, o/D Rubens/NPR Val/NPR ./, Célia/NPR Franco/NPR ./, e/CONJ um/D-UM dia/N ./, eu/PRO disse/VB-D pro/P+D rapaz/N que/WPRO arrumava/VB-D o/D cenário/N ./, disse/VB-D :/. "/QT Olha/VB-P ./, eu/PRO eu/PRO estou/ET-P com/P um/D-UM sapato/N novo/ADJ e/CONJ tu/PRO fazes/VB-P o/D seguinte/ADJ-G ./, tu/PRO pegas/N-P um/D-UM microfone/N e/CONJ botas/N-P no/P+D chão/N ./, que/C quando/CONJS eu/PRO vou/VB-P patinar/VB ./, o/D pessoal/ADJ-G pode/VB-P dar/VB a/D-F audição/N ./, né/NPR ?/.

Dos/P+D-P meus/PRO\$-P pés/N-P "/QT ./.

E/CONJ o/D cara/N :/. "/QT Tudo/Q bem/ADV "/QT ./, né/NPR ?/.

Arrumou/VB-D ali/ADV e/CONJ tal/ADJ-R-G e/CONJ coisa/N ./, né/NPR ?/.

Deixou/VB-D tudo/Q prontinho/N vou/VB-P dançar/VB e/CONJ tal/ADJ-R-G e/CONJ coisa/N ./, né/NPR ?/.

Rodando/VB-G ./, né/NPR ?/.

Programa/N ao/P+D vivo/ADJ ./, né/NPR ?/.

A/D-F guria/N de/P lá/ADV ./.

Eu/PRO entrava/VB-D ./, depois/ADV entrava/VB-D ela/PRO ./, eu/PRO me/CL lembro/VB-P que/C eu/PRO tinha/TR-D que/C dar/VB uma/D-UM-F bofetada/N nela/P+PRO ./, era/SR-D machão/N ./, ela/PRO tinha/TR-D que/WPRO apanhar/VB ./, pra/P platéia/N gostar/VB ./, né/NPR ?/.

Eu/PRO vou/VB-P fazer/VB e/CONJ ela/PRO não/NEG se/SE dá/VB-P conta/N que/WPRO eu/PRO tinha/TR-D dito/VB-PP pra/P ele/PRO botar/VB o/D microfone/N ./, e/CONJ ela/PRO vai/VB-P e/CONJ diz/VB-P :/. "/QT Olha/VB-P o/D microfone/N !/.

"/QT Esses/D-P são/SR-P os/D-P casos/N-P ./, entendeu/VB-D ?/.

Que/WPRO acontecem/VB-P ./, né/NPR ?/.

Outra/OUTRO-F vez/N foi/VB-D no/P+D Rio/NPR de/P Janeiro/NPR ./, Hilton/NPR Hotel/NPR e/CONJ todo/Q o/D artista/N ./, quase/ADV oitenta/NUM por/P cento/NUM ./, vou/VB-P dizer/VB oitenta/NUM ./, me/CL deu/VB-D um/D-UM estado/N de/P nervos/N-P ./, (( "/QT que/WPRO tu/PRO sabes/VB-P "/QT ))( uma/D-UM-F platéia/N

é/SR-P uma/D-UM-F platéia/N ,/, né/NPR ?/.

Então/ADV o/D cara/N tomou/VB-D um/D-UM traguinho/N (( "/QT vamos/VB-P pegar/VB "/QT ))( uma/D-UM-F garrafinha/N de/P uísque/NPR ,/, essas/D-F-P de/P bolso/N ,/, né/NPR ?/.

Botava/VB-D escondidinha/ADJ-F e/CONJ tal/ADJ-G ,/, né/NPR ?/.

E/CONJ chegava/VB-D no/P+D camarim/N ,/, se/SE enfeitava/VB-D ,/, botava/VB-D talco/VB-G tal/ADJ-R-G e/CONJ coisa/N ,/, por/P causa/N dos/P+D-P refletores/N-P e/CONJ aquela/D-F coisa/N toda/Q-F ,/, então/ADV dava/VB-D uma/D-UM-F bicada/VB-AN-F ,/, e/CONJ eu/PRO fui/VB-D e/CONJ dei/VB-D uma/D-UM-F bicada/VB-AN-F ,/, mas/CONJ eu/PRO estava/ET-D louco/ADJ pra/P urinar/VB ,/, mas/CONJ estava/ET-D louco/ADJ pra/P urinar/VB mesmo/ADJ ,/, e/CONJ não/NEG havia/HV-D meio/N ,/, rapaz/N ,/, de/P saber/VB onde/WADV era/SR-D o/D tal/ADJ-R-G de/P banheiro/N pra/P mim/PRO ir/VB ,/, né/NPR ?/.

Onde/WADV era/SR-D o/D WC/NPR que/C eu/PRO não/NEG sabia/VB-D nada/Q-NEG ,/, né/NPR ?/.

(( "/QT Enxerguei/VB-D "/QT ))( um/D-UM cara/N lá/ADV :/. "/QT Vem/VB-P cá/ADV ,/, onde/WADV é/SR-P "/QT O/D cara/N diz/VB-P :/. "/QT Não/NEG ,/, eu/PRO também/ADV não/NEG sei/VB-P ,/, estou/ET-P chegando/VB-G agora/ADV ,/, junto/ADV contigo/P+PRO ./.

"/QT Né/NPR ?/.

E/CONJ eu/PRO ,/, tá/ADV ,/, aí/ADV peguei/VB-P a/D-F garrafa/N mijei/VB-D ,/, puf/N fechei/VB-D e/CONJ deixei/VB-D ali/ADV ,/, né/NPR ?/.

A/D-F guria/N entra/VB-P do/P+D outro/OUTRO camarim/N ,/, vê/VB-P a/D-F garrafa/N ,/, pega/VB-P e/CONJ puf/N !/.

./.

É/SR-P fato/N que/WPRO acontece/VB-P ./.

Outra/OUTRO-F vez/N foi/SR-D aqui/ADV em/P Sã/NPR aqui/ADV em/P em/P Jaguarão/NPR ,/, nunca/ADV-NEG tinha/TR-D ido/VB-PP a/D-F Jaguarão/N ,/, não/NEG conhecia/VB-D ,/, aí/ADV tá/ADV ,/, vamos/VB-I embora/ADV pra/P Jaguarão/NPR ,/, então/ADV vamos/VB-P ,/, (( "/QT aí/ADV tá/ADV "/QT ))( ./.

O/D o/D (( "/QT Bisquintero/NPR "/QT ))( ,/, que/WPRO era/SR-D o/D empresário/N ,/, e/CONJ o/D Salim/NPR :/. "/QT Olha/VB-P ,/, vocês/PRO vão/VB-P fazer/VB um/D-UM show/N lá/ADV ,/, tá/ADV ?/.

Termina/VB-P o/D show/N ,/, pega/VB-AN-F-P a/P kombi/N ,/, se/SE manda/VB-P

que/C vocês/PRO vão/VB-P ter/TR que/C estar/ET em/P Bagé/N hem/INTJ !/.

Não/NEG esquece/VB-P ./.

"/QT "/QT Não/NEG ./, tudo/Q bem/ADV ./.

Mas/CONJ é/SR-P uma/D-UM-F apresentação/N só/FP ?/.

"/QT "/QT Não/NEG ./, é/SR-P uma/D-UM-F apresentação/N só/FP ./.

"/QT "/QT Tudo/Q bem/ADV ./, vamos/VB-I embora/ADV ./.

"/QT Aí/ADV chegou/VB-D de/P manhã/N ./, pá/N daqui/P+ADV ./, pá/N dali/P+ADV ./, pá/N dáca/NPR ./, chegou/VB-D o/D (/ (/QT Rudi/NPR "/QT )/( e/CONJ disse/VB-D pra/P mim/PRO :/. "/QT Olha/VB-P ./, Cláudio/NPR ./, a/D-F turma/N que/WPRO ia/VB-D pifou/VB-D ./, ah/INTJ !/.

Mas/CONJ o/D show/N vai/VB-P sair/VB ./, eu/PRO entrei/VB-D em/P contato/N com/P Montevideo/NPR ./, e/CONJ vão/VB-P ir/VB os/D-P músicos/N-P pra/P lá/ADV ./, vão/VB-P esperar/VB vocês/PRO lá/ADV ./.

"/QT "/QT Então/ADV tudo/Q bem/ADV ./.

"/QT Mas/CONJ eu/PRO não/NEG conhecia/VB-D nenhum/Q-NEG deles/P+PRO ./, né/NPR ?/.

Nem/CONJ-NEG sabia/VB-D quem/WPRO era/SR-D Tupamaro/NPR ./, quem/WPRO deixava/VB-D de/P ser/SR Tupamaro/NPR ./, não/NEG sabia/VB-D quem/WPRO era/SR-D ninguém/Q-NEG ./.

Aí/ADV pegamos/VB-P a/D-F guria/N ./, pegamos/VB-P esse/D rapaz/N ./, que/WPRO eu/PRO falei/VB-D ./, que/WPRO se/SE levantou/VB-D ./, né/NPR ?/.

O/D falecido/VB-AN Dudu/NPR ./, o/D apelido/N dele/P+PRO era/SR-D Dudu/NPR ./.

"/QT Dudu/NPR ./, vamos/VB-I embora/ADV ./.

"/QT "/QT Então/ADV vamos/VB-P embora/ADV ./.

"/QT Aí/ADV fomos/VB-D embora/ADV ./.

Chegamos/VB-P lá/ADV ./, uma/D-UM-F terra/N que/WPRO Deus/NPR o/CL livre/VB-SP ./, bah/INTJ !/.

Aí/ADV tá/ADV ./, descemos/VB-P na/P+D-F rodoviária/N ./, rodamos/VB-P e/CONJ rodamos/VB-P ./, aí/ADV vamos/VB-P procurar/VB um/D-UM hotel/N ./, aí/ADV o/D rapaz/N que/WPRO estava/ET-D na/P+D-F ro/N na/P+D-F rodoviária/N nos/P+D-P esperando/VB-G ./, chegou/VB-D pra/P nós/PRO :/. "/QT Olha/VB-P ./, tem/TR-P uma/D-UM-F zebra/N pra/P vocês/PRO ./.

"/QT Digo/VB-P :/. "/QT Aí/ADV ./, pára/VB-P aí/ADV ./, outra/OUTRO-F ./.

Qual/WPRO é/SR-P ?/.



"/QT "/QT Os/D-P hotéis/N-P estão/ET-P tudo/Q lotados/VB-AN-P ./.

"/QT Né/NPR ?/.

E/CONJ eu/PRO digo/VB-P :/. "/QT Mas/CONJ e/CONJ agora/ADV ./, meu/PRO\$ filho/N ?/.

"/QT "/QT Não/NEG te/CL preocupas/VB-P ./, vamos/VB-P dar/VB um/D-UM jeito/N (( "/QT nós/PRO vamos/VB-P arrumar/VB umas/D-UM-F-P dessas/P+D-F-P pensões/N-P aí/ADV "/QT ))( ./, dormitório/N dos/P+D-P caminhoneiros/N-P e/CONJ que/WPRO é/SR-P pra/P quebrar/VB o/D galho/N só/FP ./, né/NPR ?/.

Pra/P descansar/VB aí/ADV um/D-UM pouco/Q e/CONJ tal/ADJ-R-G e/CONJ coisa/N ./, né/NPR ?/.

Depois/ADV nós/PRO vamos/VB-P seguir/VB viagem/N ./.

"/QT "/QT Ah/INTJ !/.

Tudo/Q bem/ADV ./.

Não/NEG tem/TR-P problema/N ./, vamos/VB-I embora/ADV ./.

"/QT Aí/ADV tá/ADV ./, mas/CONJ a/D-F guria/N nervosa/ADJ-F que/WPRO estava/ET-D ./, sabe/VB-P ?/.

A/D-F guria/N novinha/ADJ-F ./, nervosa/ADJ-F ./, aí/ADV :/. "/QT Tá/NPR ./, vamos/VB-P pra/P lá/ADV ./.

"/QT Aí/ADV fomos/VB-D ./, chegamos/VB-D lá/ADV na/P+D-F pensão/N lá/ADV ./.

A/D-F senhora/NPR prontamente/ADJ-G arrumou/VB-D um/D-UM quarto/N de/P casal/N e/CONJ um/D-UM quarto/N de/P solteiro/ADJ pra/P ele/PRO ./.

Nem/CONJ perguntou/VB-D se/WQ éramos/SR-D casados/VB-AN-P ou/CONJ se/CONJS não/NEG éramos/SR-D ./, né/NPR ?/.

Viu/VB-D é/SR-P casado/VB-AN ./.

Então/ADV tá/ADV ./, tudo/Q bem/ADV ./, fomos/VB-D pra/P lá/ADV ./, aí/ADV tá/ADV (( "/QT certo/N "/QT ))( ficamos/VB-P ali/ADV ./, se/SE arrumamos/VB-P ./, trocamos/VB-P de/P roupa/N ./, aquela/D-F coisa/N toda/Q-F ./, tomamos/VB-P banho/N ./.

"/QT Olha/VB-P ./, vamos/VB-P jantar/VB ./, né/NPR ?/.

Vamos/VB-P se/SE preparar/VB pra/P jantar/N e/CONJ coisa/N e/CONJ vamos/VB-P até/P o/D clube/N ./, pra/P conhecer/VB o/D clube/N ./.

"/QT O/D show/VB-AN era/SR-D à/P+D-F meianoite/N ./, né/NPR ?/.

Primeiro/ADV entrava/VB-D o/D conjunto/N pra/P depois/ADV entrar/VB a/D-F típica/N ./, a/D-F típica/ADJ-F sempre/ADV é/SR-P é/SR-P a/D-F última/ADJ-F ./,

né/NPR ?/.

Aí/ADV tá/ADV ./, aí/ADV fomos/VB-D pro/P+D restaurante/N ./, mas/CONJ um/D-UM restaurante/N lindo/ADJ ./, dois/NUM pisos/N-P ./.

Chegamos/VB-P no/P+D restaurante/N e/CONJ era/SR-D uma/D-UM-F gurizada/N ./, mas/CONJ gurizada/VB-AN-F ./, mas/CONJ era/SR-D gurizada/VB-AN-F mesmo/FP ./, sabe/VB-P o/D que/WPRO que/WPRO era/SR-D ?/.

Tudo/Q :/. pivetezinho/N ./, mas/CONJ gurizada/N mesmo/ADJ ./.

Isso/DEM eu/PRO nunca/ADV-NEG me/CL esqueço/VB-P na/P+D-F minha/PRO\$-F vida/N ./, nunca/ADV vou/VB-P esquecer/VB também/ADV ./, mas/CONJ tudo/Q gurizada/VB-AN-F ./, era/SR-D tomando/VB-G CocaCola/NPR ./, tomando/VB-G cerveja/N (( "/QT essa/D-F coisa/N "/QT ))( e/CONJ nós/PRO entramos/VB-P ./, fomos/SR-D lá/ADV pra/P cima/ADV ./, jantarmos/VB-F ./.

Aí/ADV tá/ADV ./, estamos/ET-P na/P+D-F mesa/N ali/ADV olha/VB-P pra/P mim/PRO ./, ele/PRO olha/VB-P pra/P mim/PRO ./, o/D meu/PRO\$ companheiro/N esse/D ./, e/CONJ diz/VB-P :/. "/QT Bah/NPR !/.

Cláudio/NPR ./, ah/INTJ !/.

Vamos/VB-I embora/ADV sabe/VB-P que/C tu/PRO já/ADV pensou/VB-D ./, tu/PRO já/ADV ouviu/VB-D quando/CONJS tu/PRO entrou/VB-D esses/D-P guris/N-P falando/VB-G tudo/Q que/WPRO já/ADV estão/ET-P com/P o/D convite/N na/P+D-F mão/N pra/P ir/VB pro/P+D baile/N ./, e/CONJ vocês/PRO dançar/VB tango/N e/CONJ esses/D-P guris/N-P aí/ADV só/FP têm/TR-P agora/ADV o/D iêiêiê/VB deles/P+PRO ./.

Bah/NPR !/.

Vai/VB-P ser/SR horrível/ADJ-G ./.

"/QT Aí/ADV ./, a/D-F guria/N tremia/VB-D ./.

"/QT Vamos/VB-P firme/ADJ-G ./, né/NPR ?/.

"/QT Aí/ADV ./, tá/ADV ./, sentamos/VB-D numa/P+D-UM-F mesa/N e/CONJ coisa/N ./, pedimos/VB-P a/P janta/VB-P ./, aí/ADV nesse/P+D meio/ADJ tempo/N nós/PRO ficamos/VB-P jantando/VB-G ./, aí/ADV chegou/VB-D um/D-UM um/D-UM rapaz/N que/WPRO é/SR-P ./, casualmente/ADV é/SR-P gerente/ADJ-G :/. da/P+D-F Caixa/NPR Econômica/NPR Federal/NPR e/CONJ se/SE dava/VB-D comigo/P+PRO aqui/ADV no/P+D Centro/NPR ./, foi/VB-D companheiro/N meu/PRO\$ aqui/ADV do/P+D Centro/NPR também/ADV ./, aí/ADV quando/CONJS "/QT Olha/VB-P ./, Cláudio/NPR ./, tal/ADJ-R-G e/CONJ coisa/N e/CONJ bábábá/N ./, o/D que/WPRO

está/ET-P fazendo/VB-G aqui/ADV ?/.

"/QT Aí/ADV contei/VB-D pra/P ele/PRO ,/, né/NPR ?/.

Digo/VB-P :/. "/QT Bah/NPR !/.

Eu/PRO estou/ET-P fazendo/VB-G isso/DEM ./.

"/QT "/QT O/D quê/WPRO ?/.

Não/NEG pode/VB-P ser/SR !/.

"/QT "/QT É/SR-P ./.

"/QT (( "/QT Disse/VB-D "/QT ))( :/. "/QT Bah/NPR !/.

Mas/CONJ não/NEG é/SR-P possível/ADJ-G ,/, é/SR-P tu/PRO o/D bailarino/N ,/, bah/INTJ !/.

Mas/CONJ tu/PRO não/NEG queiras/N-P saber/VB ,/, rapaz/N ,/, bah/INTJ !/.

Isso/DEM aqui/ADV já/ADV está/ET-P tudo/Q espalhado/VB-AN tá/ADV ,/, mas/CONJ espalhou/VB-D mesmo/ADJ ,/, diz/VB-P que/C vem/VB-P um/D-UM casal/N de/P bailarinos/N-P fora/ADV de/P série/N ,/, propaganda/N aí/ADV ,/, bah/INTJ !/.

Mas/CONJ tem/TR-P assim/ADV que/C vou/VB-P te/CL contar/VB ,/, né/NPR ?/.

Tu/PRO vais/VB-P no/P+D clube/N daqui/P+ADV ah/INTJ !/.

Tu/PRO vais/VB-P ver/VB que/C tem/TR-P de/P propaganda/N !/.

"/QT Digo/VB-P :/. "/QT Bah/NPR !/.

"/QT E/CONJ a/D-F guria/ADJ-F tremia/VB-D como/CONJS uma/D-UM-F capivara/N ,/, aí/ADV tudo/Q bom/ADJ ,/, aí/ADV jantamos/VB-D tinha/TR-D um/D-UM guia/N nosso/PRO\$ ,/, né/NPR ?/.

Que/C era/SR-D empregado/VB-AN do/P+D em/P do/P+D empresário/N ,/, aí/ADV fomos/VB-D pra/P o/D clube/N ,/, bah/INTJ !/.

E/CONJ chegamos/VB-D no/P+D clube/N parede/N e/CONJ tudo/Q e/CONJ cheia/ADJ-F de/P cartaz/VB e/CONJ bábábá/VB ,/, aí/ADV a/D-F guria/N mais/ADV-R tremia/VB-D ainda/ADV ,/, aí/ADV entramos/VB-P pra/P uma/D-UM-F sala/N lá/ADV da/P+D-F diretoria/N ,/, não/NEG tinha/TR-D camarim/N ,/, não/NEG tinha/TR-D nada/Q-NEG não/NEG tem/TR-P mesmo/ADJ ,/, aí/ADV entramos/VB-P ,/, estavam/ET-D os/D-P caras/N-P ali/ADV ,/, a/P arrumando/VB-G os/D-P aparelhos/N-P deles/P+PRO ,/, né/NPR ?/.

"/QT Oi/N ,/, oi/N ,/, oi/N ./.

Ah/INTJ !/.

Sou/SR-P brasileiro/ADJ ./.

"/QT "/QT Ah/INTJ !/.

É/SR-P brasileiro/ADJ ?/.

"/QT "/QT Sou/SR-P brasileiro/ADJ ./.

"/QT "/QT (( "/QT Usted/NPR "/QT ))( é/SR-P bailarino/N de/P tango/N ?/.

Por/P favor/N ./.

"/QT "/QT Sim/ADV porquê/WADV ?/.

"/" Bom/ADJ ./, aí/ADV ele/PRO ficou/VB-D meio/ADJ assim/ADV ./, né/NPR ?/.

O/D castelhano/N e/CONJ ela/PRO :/. "/QT Ai/INTJ ./, ai/INTJ ./, ai/INTJ Cláudio/NPR ./, vamos/VB-I embora/ADV ./, vamos/VB-P se/SE mandar/VB ./.

"/QT Então/ADV ela/PRO tinha/TR-D ela/PRO mudava/VB-D de/P roupa/N ./, dois/NUM trajes/N-P ./, né/NPR ?/.

Depois/ADV outra/OUTRO-F ./, aí/ADV ./, fomos/VB-D na/P+D-F mala/N e/CONJ cadê/VB a/D-F roupa/N ?/.

Faltava/VB-D a/D-F saia/N ./, a/D-F blusa/N tinha/TR-D ./, mas/CONJ a/D-F saia/N não/NEG tinha/TR-D ./, a/D-F saia/N era/SR-D rasgada/VB-AN-F até/P aqui/ADV ./, quer/VB-P dizer/VB ./, não/NEG era/SR-D rasgada/VB-AN-F ./, (( "/QT era/SR-D costurada/VB-AN-F ./, né/NPR ?/.

"/QT ))( Era/SR-D rasgada/VB-AN-F até/P aqui/ADV ./.

Cadê/VB a/D-F saia/N ?/.

Não/NEG tinha/TR-D ./.

E/CONJ agora/ADV tudo/Q ./, né/NPR ?/.

Conforme/CONJS for/SR-SR vamos/VB-P ./.

Aí/ADV os/D-P caras/N-P estão/ET-P lá/ADV ./, com/P os/D-P instrumentos/N-P ./, tocando/VB-G e/CONJ coisa/N ./, né/NPR ?/.

Afinando/VB-G os/D-P instrumentos/N-P deles/P+PRO ./, aí/ADV o/D cara/N larga/ADJ-F uma/D-UM-F música/N lá/ADV e/CONJ eu/PRO pá/N !/.

Gritei/VB-D o/D nome/N da/P+D-F música/N e/CONJ eu/PRO pá/N !/.

Daqui/P+ADV a/P pouco/Q se/SE levantou/VB-D um/D-UM Tupamaro/NPR ./, mas/CONJ era/SR-D um/D-UM baita/N Tupamaro/NPR :/. "/QT O/D !/.

Muchacho/N yo/ADJ soy/VB-D el/D cantor/N ./, usted/VB-D no/P+D é/SR-P nada/Q-NEG cá/ADV "/QT ./, ./.

"/QT Aí/ADV ./, nós/PRO tivemos/TR-D que/C sair/VB correndo/VB-G ./, né/NPR ?/.

Aí/ADV saímos/VB-D ./.

A/D-F guria/N disparou/VB-D ./.

Fui/VB-D encontrar/VB ela/PRO na/P+D-F pensão/N ./, que/C a/D-F pensão/N ficava/VB-D umas/D-UM-F-P quatro/NUM quadras/N-P (( "/QT numa/P+D-UM-F "/QT ))( rua/N reta/ADJ-F lá/ADV ./.

Era/SR-D a/D-F única/ADJ-F que/WPRO ta/CL+CL que/C era/SR-D calçada/N ./, também/ADV ./, né/NPR ?/.

Aí/ADV fui/VB-D lá/ADV :/. "/QT Vamos/VB-P ./, que/WPRO é/SR-P isso/DEM ?/.

"/QT "/QT Não/NEG ./, mas/CONJ está/ET-P louco/ADJ ./, tu/PRO vais/VB-P brigar/VB com/P o/D homem/N ./.

"/QT "/QT Não/NEG ./, que/C nada/Q-NEG ./, vamos/VB-P com/P calma/N ./.

"/QT "/QT Não/NEG ./, mas/CONJ vai/VB-P dar/VB zebra/N ./.

"/QT "/QT Não/NEG vai/VB-P dar/VB zebra/N ./, vamos/VB-I embora/ADV ./.

"/QT Aí/ADV tá/ADV ./, aí/ADV se/SE arrumamos/VB-P tal/ADJ-R-G e/CONJ coisa/N ./, fomos/VB-D pra/P lá/ADV ./, levamos/VB-P as/D-F-P malas/N-P ./, as/D-F-P roupas/N-P ./, né/NPR ?/.

Que/WPRO tinha/TR-D que/WPRO levar/VB ./, aí/ADV aí/ADV cheguei/VB-D pro/P+D cara/N e/CONJ disse/VB-D :/. "/QT Olha/VB-P ./, o/D negócio/N é/SR-P o/D seguinte/ADJ-G "/QT :/. tinha/TR-D dois/NUM que/WPRO tocavam/VB-D bandonion/N ./, disse/VB-D :/. "/QT Olha/VB-P ./, o/D problema/N é/SR-P o/D seguinte/ADJ-G ./, quando/CONJS eu/PRO fizer/VB-SR com/P o/D dedo/N pra/P cima/ADV tu/PRO rasgas/N-P o/D bandonion/N ./, quer/VB-P dizer/VB ./, eu/PRO estou/ET-P ouvindo/VB-G a/D-F música/N e/CONJ sei/VB-P o/D exato/ADJ momento/N ./, né/NPR ?/.

Que/C ele/PRO vai/VB-P fazer/VB o/D babado/VB-AN ./.

Aí/ADV tá/ADV ./, tudo/Q bem/ADV ./, aí/ADV tá/ADV tal/ADJ-G e/CONJ coisa/N fui/SR-D lá/ADV ./, entramos/VB-P no/P+D salão/N ./.

Bah/NPR !/.

Todo/Q mundo/N ./, e/CONJ aplaudiu/VB-D ./, né/NPR ?/.

Nessas/P+D-F-P alturas/N-P estava/ET-D tudo/Q bem/ADV ./, e/CONJ as/D-F-P luzes/N-P ficaram/VB-D meia/NUM foscas/N-P ./, e/CONJ nós/PRO saímos/VB-D dançando/VB-G ./, um/D-UM baita/N de/P um/D-UM salão/N ./, todo/Q mundo/N sentadinho/N bonito/ADJ ./, sabe/VB-P ?/.

Coisa/N mais/ADV-R linda/ADJ-F ./, todo/Q mundo/N sentado/VB-AN ./, aí/ADV tá/ADV saímos/VB-D dançando/VB-G e/CONJ coisa/N ./, daqui/P+ADV a/P pouco/Q eu/PRO vou/VB-P ./, dou/VB-P uma/D-UM-F volta/N nela/P+PRO e/CONJ pim/N !/.

Um/D-UM tapa/VB-P na/P+D-F bunda/N dela/P+PRO ./, desculpe/VB-SP o/D termo/N ./, né/NPR ?/.

Mas/CONJ pim/N !/.

Tapa/VB-P na/P+D-F bunda/N ./, daí/P+ADV quando/CONJS eu/PRO dou/VB-P o/CL tapa/VB-P "/QT Dálhe/NPR machão/NPR !/.

"/QT Aí/ADV o/D pessoal/ADJ-G ./, entendeu/VB-D ?/.

Ficou/VB-D mais/ADV-R à/P+D-F vontade/N ./, aí/ADV tá/ADV ./, deu/VB-D tal/ADJ-R-G e/CONJ coisa/N e/CONJ eu/PRO fiz/VB-D sinal/N pro/P+D cara/N ./, o/D cara/N rasga/VB-P e/CONJ eu/PRO ./, deito/VB-P ela/PRO ./, né/NPR ?/.

Na/P+D-F perna/N ./, né/NPR ?/.

Ah/INTJ !/.

E/CONJ o/D pessoal/ADJ-G sai/VB-P aplaudindo/VB-G ./.

Terminou/VB-D e/CONJ "/QT Vamos/VB-P sair/VB pro/P+D corredor/N "/QT ./, digo/VB-P :/. "/QT vamos/VB-P sair/VB ./.

"/QT Mas/CONJ o/D pessoal/ADJ-G de/P cima/ADV de/P invés/N de/P ficar/VB sentadinho/N ./, e/CONJ nós/PRO não/NEG estávamos/ET-D prestando/VB-G atenção/N ./, vieram/VB-D todos/Q-P pra/P copa/N ./, mas/CONJ toda/Q-F a/D-F gurizada/N ./, aí/ADV no/P+D aí/ADV a/D-F gurizada/N queria/VB-D meu/PRO\$ chapéu/N ./, queria/VB-D a/D-F minha/PRO\$-F roupa/N ./, queria/VB-D um/D-UM pedaço/N de/P um/D-UM um/D-UM pedaço/N de/P outra/OUTRO-F coisa/N ./, mas/CONJ nós/PRO saímos/VB-D apavorados/VB-AN-P ./, escoltados/VB-AN-P até/P ./.

## 8.1 Etiquetas

TYPE	TAG	VALUE	EXAMPLES
SER (to be - individual level))	SR	Infinitive (in verbal and nominal contexts)	a ser/SR exemplo de emendas
	SR-F	Inflected Infinitive (with explicit morphological marks)	para sermos/SR-F completos e felices
	SR-I	Imperative (with explicit morphological marks)	Sê/SR-I fiel a Deus, e ele sempre te amará.
	SR-P	Present	não é/SR-P muito que assim me acolha...
	SR-SP	Present Subjunctive	que eles sejam/SR-SP tão teimosos... seja/SR-SP Deus louvado...
	SR-D	Past	Fatalidade foi/SR-D faltar-vos carta minha ... Boa maré era/SR-D esta ...
	SR-RA	Verbal forms with the inflectional morpheme -ra	Past Perfect Ele fôra/SR-RA embora,...  Conditional  Hypothetical Se fôra/SR-RA Deus, como vós sois,...
	SR-SD	Past Subjunctive	...que fosse/SR-SD contra a minha regra...
	SR-R	Future and Conditional	esta lembrança será/SR-R, mais que necessidade, imprudência ...  seria/SR-R possível que ...
	SR-SR	Future Subjunctive	mais ditoso que Quando fordes/SR-SR rei
	SR-G	Gerund	sendo/SR-G tão má cousa como sou
SR-PP	Perfect Participle	como haver sido/SR-PP tão má	

TYPE	TAG	VALUE	EXAMPLES
HAYER (there to be)	HV	Infinitive (in verbal and nominal contexts)	em tudo isto pode haver/HV vaidade

	HV-F	Inflected Infinitive (with explicit morphological marks)	quão lastimosa coisa será haverem/HV-F de deixar aquelas almas
	HV-I	Imperative (with explicit morphological marks)	Havei/HV-I vós....
	HV-P	Present	em toda parte há/HV-P sido servido fazer-me
	HV-SP	Present Subjunctive	...fingir que tem reparado em que eu haja/HV-SP faltado assim.
	HV-D	Past	...o amor que vos tem, havia/HV-D de furtar-lho.
	HV-RA	Verbal forms with the inflectional morpheme -ra	Past Perfect Conditional Hypothetical (...), como se houvera/HV-RA trabalhado.
	HV-SD	Past Subjunctive	Não cuidei que houvesse/HV-SD tão pouco coração.
	HV-R	Future and Conditional	...o que haverá/HV-R em contrário. Haveria problemas, caso.....
	HV-SR	Future Subjunctive	Se me houver/HV-SR perdoado as culpas...
	HV-G	Gerund	Havendo/HV-G provimentos, só os envie se...
	HV-PP	Perfect Participle	...tem havido/HV-PP nesse convento.
	HV-NA	Passive Participle ( <a href="#">see agreement</a> and <a href="#">participles as adjectives</a> )	As casas havidas/HV-AN-F-P

TYPE	TAG	VALUE	EXAMPLES
ESTAR (to be - stage level)	ET	Infinitive (in verbal and nominal contexts)	é estar/ET nas mãos de meu Senhor Jesu Cristo,...
	ET-F	Inflected Infinitive (with explicit morphological marks)	Assim, para estarmos/ET-F com Nosso Senhor,...
	ET-I	Imperative (with explicit morphological marks)	Estai/ET-I agora junto ao amor de...
	ET-P	Present	que estão/ET-P no côro.
	ET-SP	Present Subjunctive	não esteja/ET-SP como eu



			cuidava.
	ET-D	Past	que já estavam/ET-D de espera...
	ET-RA	Verbal forms with the inflectional morpheme -ra	Past Perfect Bem aviado estivera/ET-RA eu...  Conditional  Hypothetical Se estivera/ET-RA
	ET-SD	Past Subjunctive	Espero em Deus que nunca estivesse/ET-SD azedo...
	ET-R	Future and Conditional	que cêdo estará/ET-R em Deus...  que cêdo estaria/ET-R.....
	ET-SR	Future Subjunctive	Quando estiveres/ET-SR melhor....
	ET-G	Gerund	Estando/ET-G inocente...
	ET-PP	Perfect Participle	Tinha estado/ET-PP em Mafra..

TYPE	TAG	VALUE	EXAMPLES
TER (to have)	TR	Infinitive (in verbal and nominal contexts)	a ter/TR conta o que se passara..
	TR-F	Inflected Infinitive (with explicit morphological marks)	de terem/TR-F as noviças...
	TR-I	Imperative (with explicit morphological marks)	Tende/TR-I esperança...
	TR-P	Present	Tenho/TR-P a Madre de Deus por mim...
	TR-SP	Present Subjunctive	Não tenhais/TR-SP cuidado...
	TR-D	Past	Que teve/TR-D seu próprio filho...
	TR-RA	Verbal forms with the inflectional morpheme -ra	Past Perfect Tivera/TR-RA bondade e simplicidade...  Conditional  Hypothetical Se tivera/TR-RA eu....
	TR-SD	Past Subjunctive	Se tivesse/TR-SD desejo...

	TR-R	Future	Como o terá/TR-R para fazer um sermão
	TR-SR	Future Subjunctive	Se Deus não me tiver/TR-SR muito...
	TR-G	Gerund	Tendo/TR-G sido encontrado...
	TR-PP	Perfect Participle	Tido/TR-PP paciência e afeto...
	TR-NA	Passive Participle ( <a href="#">see agreement</a> and <a href="#">participles as adjectives</a> )	Tidas/TR-AN-F-P as cartas...

TYPETAGVALUE	EXAMPLES		
VERBS in general	VB	Infinitive (in verbal and nominal contexts)	...e bastará ao menos dar/VB três vezes Ordens no ano, ...  ...é a que retrata com mais semelhança o falar/VB e conversação dentre os amigos  O amar/VB a Deus (see <a href="#">infinitives as nouns</a> )
	VB-F	Inflected Infinitive (with explicit morphological marks)	Segue-se agora tratarmos/VB-F de como uma pessoa se alongará...
	VB-I	Imperative (with explicit morphological marks)	Sabei/VB-I que agora estais no Cabo da Boa Esperança...
	VB-P	Present	Bem me parece/VB-P isto, porque...  Creio/VB-P em Jesu Cristo,...
	VB-SP	Present Subjunctive	..., até que nos vejamos/VB-SP
	VB-D	Past	Só vos digo que estimei/VB-D as vossas novas....
	VB-RA	Verbal forms with the inflectional morpheme -ra	Past Perfect Bem quisera/VB-RA eu prestar para vo-lo agradecer, ...  Conditional E se mais mundo houvera, lá chegara/VB-RA  Hypothetical Se ninguém no mundo se

			salvara/VB-RA, então (...)
	VB-SD	Past Subjunctive	Se quisesse/VB-SD meu Deus que....
	VB-R	Future	...de quem não verei/VB-R por muitos anos
	VB-SR	Future Subjunctive	Quando eu puder/VB-SR lhe darei algum....
	VB-G	Gerund	Sabendo/VB-G que...
	VB-PP	Perfect Participle	Terá feito/VB-PP os trabalhos...
	VB-AN	Passive Participle ( <a href="#">see agreement</a> and <a href="#">adjective as participles</a> )	Foi servido/VB-AN de reduzir ao seu rebanho esta ovelha perdida .... Aparecera Cristo crucificado/VB-AN

TYPE	TAG	VALUE	APPLICATION	EXAMPLES
GENDER	none	Masculine	Default	que o/D dia/N
	-F	Feminine	Determiners	que a/D-F noite
			Adjectives	a noite fermosa/ADJ-F
			Possessive Pronouns	de sua/PRO\$-F fermosura
			Quantifiers	de toda/Q-F a alma
			Passive Participles	a ser feita/VB-AN-F
			Relative Pronoun 'cuja'	de cuja/WPRO\$-F virtude
			Cardinal Number 'duas'	de duas/NUM-F fragatas
	-G	Double gender	Determiner 'tal'	Tal/D-G saída não nos perturbara.
			Adjectives	da grande/ADJ-G revolta de...
Quantifiers			de cada/Q-G homem...	
NUMBER	None	Singular	Default	que o/D dia/N
	-P	Plural	Nouns	as terras/N-P...

			Determiners	Os/D-P mares e as/D-F-P terras....
			Adjectives	Dois fermosos/ADJ-P frutos... Duas fermosas/ADJ-F-P frutas..
			Possessive Pronouns	Seus/PRO\$-P favores e....
			Quantifiers	Muitos/Q-P louvores...
			Passive Participles	Foram encontrados/VB-AN-P mulheres e crianças....
			Some Relative Pronouns	As correntes cujos/WPRO\$-P elos....

TYPE	TAG	VALUE	APPLICATION	EXAMPLES
NOUNS	N	Singular	Common nouns	A gente/N de Lisboa Concordo com este parecer/N Este criado/N de Vossa Senhoria contraria ao bem/N commum
			Others	O contrário/N é verdadeiro Sobre este particular/N, penso que
			Collective nouns	A multidão/N
			Infinitive forms largelly recognised as nouns	O jantar/N fôra agradabilíssimo <i>versus</i> <b>Jantar/VB com amigos é algo como...</b>
	N-P	Plural	Common nouns	Mares/N-P e espumas/N-P Jornais/N-P, meu caro

				amigo...
PROPER NOUNS	NPR	Singular	Common nouns with initial capital letter  (except after a period punctuation – see above)	... leva na mão direita o Ceptro/NPR , com que ha de governar o Povo/NPR em Paz/NPR .
			<i>All Address Forms</i> , even those with non initial capital letter	Não diga Vossa/PRO\$-F Senhoria/NPR.....  Por isso, dê a senhora/NPR graças a Nosso Senhor, pois...
			Titles  and  Epithets	Ilustríssimo/ADJ-S Senhor/NPR Gonçalves/NPR de/P Azevedo/NPR  Reverendíssima/ADJ-S-F Madre/NPR  Padre/NPR Provincial/ADJ-G <i>and</i> O Provincial/ADJ-G informa que...  Espírito/NPR Santo/ADJ  Frei/NPR António/NPR  Dom/NPR Frederico/NPR  Ao ver El-Rei/NPR ao longe....  Por Monsieur/NPR Estrada/NPR  o Rey/NPR Catholico/NPR
			Dates, days of week, months, and holidays	Sábado/NPR e domingo/NPR  Dezembro/NPR e março/NPR  Dia/NPR de São/NPR João/NPR

				O dia/N de Natal/NPR Sexta/NPR Feira/NPR da Paixão/NPR A Paixão/NPR de Cristo/NPR
			Places	O/D Amazonas/NPR Portugal/NPR Rio/NPR de/P Janeiro/NPR a rua/N Nova/NPR
			Other Proper Nouns	São/NPR Bertolameu/NPR João/NPR Alves/NPR
	NPR-P	Plural	Address Forms	Suas/PRO\$-F-P Majestades/NPR-P e Altezas/NPR-P
			Titles	As Condessas/NPR-P de Braga
			Other Proper Nouns	Os/D-P Mendonças/NPR-P
TYPE	TAG	VALUE/APPLICATION	EXAMPLES	
PRONOUNS	PRO	Personal pronouns	Eu/PRO, tu/PRO, ele/PRO, ela/PRO, você/PRO, nós/PRO, vós/PRO, eles/PRO, elas/PRO, vocês/PRO	
		Oblique forms	Mim/PRO, ti/PRO, si/PRO	
	P+PRO	Preposition +oblique forms	Comigo/P+PRO, Contigo/P+PRO, Consigo/P+PRO, Conosco/P+PRO, Convosco/P+PRO	
POSSESSIVE PRONOUNS	PRO\$	Masculine singular	Meu/PRO\$, teu/PRO\$, seu/PRO\$, nosso/PRO\$, vosso/PRO\$	
	PRO\$-F	Feminine singular	Minha/PRO\$-F, tua/PRO\$-F, sua/PRO\$-F, nossa/PRO\$-F, vossa/PRO\$-F	
	PRO\$-P	Masculine plural	Meus/PRO\$-P etc...	
	PRO\$-F-P	Feminine plural	Minhas/PRO\$-F-P etc...	

CLITICS	CL	Clitics in general <i>me, te, o(s), a(s), lhe(s), nos, vos</i>	Já o/CL havia dado a mi.... Quando lhe/CL visse.... Dar-te-ei/VB-R!CL Entregou-me/VB-D+CL antes de...
	SE	Clitic "se" in all its functions	porque todos para consigo são como os olhos, que vendo tudo, não se/SE vem a si mesmos;
	CL+CL	Clitics contraction in general	e não ma/CL+CL deu

TYPE	TAG	EXAMPLES
verb+one clitic	...+CL	haverem-nos/HV-F+CL, terem-na/TR-F+CL, mostrou-o/VB-D+CL, tergiversandolhe/VB-G+CL,
verb+two clitics	...+CL+CL	permittindo-lho/VB-G+CL+CL, quadruplicandolho/VB-G+CL+CL

TAG	EXAMPLES
SR-R!CL	ser-lhe-há/ SR-R!CL
ET-R!CL	estar-me-iam/ ET-R!CL
HV-R!CL	haver-te-ei/ HV-R!CL
TR-R!CL	ter-lhe-ia/ TR-R!CL
VB-R!CL	amar-te-ei/ VB-R!CL

TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
DEFINITE DETERMINERS AND INFLECTED DEMONSTRATIVES	D	Masculine singular	O/D homem Este/D homem Aquele/D homem e o mandou para este/D.
	D-F	Feminine singular	A/D-F videira (see 'a' as <a href="#">preposition</a> ) Esta/D-F videira Aquela/D-F videira
	D-P	Masculine plural	Os/D-P homens Estes/D-P homens Aqueles/D-P homens

	D-F-P	Feminine plural	As/D-F-P videiras Estas/D-F-P videiras Aqueles/D-F-P videiras		
	D-G	Double gender singular	Tal/D-G homem  porque em tal/D-G parte vale muito  Tal/D-G videira  (see 'tal' as <a href="#">other demonstratives</a> , as <a href="#">comparative adjective</a> , and <a href="#">comparative adverb</a> )		
	D-G-P	Double gender plural	Tais/D-G-P homens Tais/D-G-P videiras		
INDEFINITE DETERMINERS/  CARDINAL NUMBER 'UM'	D-UM	Masculine singular	Um/D-UM homem  Um/D-UM quilo de forragem		
	D-UM-F	Feminine singular	Uma/D-UM-F flor  Uma/D-UM-F dúzia de ovos...		
	D-UM-P	Masculine plural	Uns/D-UM-P homens  Uns/D-UM-P quilos de forragem  Acabava por não saber como agradá-los: a uns/D-UM-P oferecia amor, a outros dedicação.....		
	D-UM-F-P	Feminine plural	Umas/D-UM-F-P flores  Umas/D-UM-F-P dúzias		
OTHER DEMONSTRATIVES	DEM	Invariable demonstratives	Por isso/DEM a razão derrubou os ídolos...  Sobre aquilo/DEM do convento, cêdo nos veremos e então falaremos.		
		'tal'	Apenas por tal/DEM faria a guerra (see 'tal' as <a href="#">comparative adjectives</a> )		
TYPE		TAG	VALUE	APPLICATION	EXAMPLES



ADJECTIVES	ADJ	Masculine singular	General adjectives	Por ser bonito/ADJ Um e meio/ADJ por cento Ano e meio/ADJ Meio/ADJ por cento
			'mesmo' as adjective	O mesmo/ADJ homem foi encontrado... O mesmo/ADJ se pode dizer... (see 'mesmo' as as focus particle, and as member of complex subordinating conjunction)
			Ordinal numbers (see <a href="#">cardinal numbers</a> )	O segundo/ADJ homem que partiu
	ADJ-F	Feminine singular		Uma ferrosa/ADJ-F igreja ...
	ADJ-G	Double-gender singular	General adjectives	Um homen grande/ADJ-G e uma mulher amável/ADJ-G
			Others	Por conseguinte/ADJ-G, ... (see <a href="#">adverbs as prepositional phrases</a> ) <b>NEW</b> não/NEG obstante/ADJ-G (see <a href="#">clusters</a> )
	ADJ-P	Masculine plural		Homens bonitos/ADJ-P
	ADJ-F-P	Feminine plural		Mulheres bonitas/ADJ-F-P

			General adjectives	Homens e mulheres valentes/ADJ-G-P
	ADJ-G-P	Double-gender plural	'demais'	Os demais/ADJ-G-P homens saíram. As demais/ADJ-G-P foram entregues...
EXCLAMATIVE/COMPARATIVE ADJECTIVES	ADJ-R	Masculine singular ('tanto')		E aplicou tanto/ADJ-R esforço <b>NEW</b> <u>que/C</u> impediu o levantar da barreira, onde....
	ADJ-R-F	Feminine singular ('tanta')		A vontade foi tanta/ADJ-R-F <b>NEW</b> <u>que/C</u> acabou por .....
	ADJ-R-P	Masculine plural ('tantos')		...do coração que há tantos/ADJ-R-P tempos vive oculto nos mistérios do recato. E aplicou tantos/ADJ-R-P esforços <b>NEW</b> <u>que/C</u> impediu o levantar da barreira, onde....
	ADJ-R-F-P	Feminine plural ('tantas')		..., pois êle foi servido de reduzir ao seu rebanho esta ovelha perdida, quási no mesmo tempo que entre as trevas do pecado e entre os horrores de tantas/ADJ-R-F-P culpas se aumentavam os descaminhos e creciam os despenhadeiros.
	ADJ-R-G	'tal'		Muito estimara eu que fôsse tal/ADJ-R-G o meu desembaraço, <b>NEW</b> <u>que/C</u> pudesse escrever a Vossa Mercê mais a miúdo. Fez um tal/ADJ-R-G esforço <b>NEW</b> <u>que/C</u> não pode....
	ADJ-R-G	'maior'/'menor'		Este é o maior/ADJ-R-G bem de todos... o grande amor que vos tem é maior/ADJ-R-G do/P+D que/WPRO cuidais.

		'melhor'/'pior'	<p>desejando de empregar em melhor/ADJ-R-G guerra o que me ficava de vida.</p> <p>E isto acabou por ser melhor/ADJ-R-G do/P+D que/WPRO descansar,....</p> <p>uma forma melhor/ADJ-R-G <b>NEW</b> <u>que/C</u> esta</p> <p>Este é o pior/ADJ-R-G ....</p> <p>O melhor/ADJ-R-G a fazer é isto.....</p>
	ADJ-R-G-P	'tais'	Esses tais/ADJ-R-G-P fatos acabaram por trazer....
		'maiores'/'menores'	...projetos maiores/ADJ-R-G-P <b>NEW</b> <u>que/C</u> aqueles....
		'melhores'/'piores'	...melhores/ADJ-R-G-P casas <b>NEW</b> <u>que/C</u> essas
SUPERLATIVE ADJECTIVES	ADJ-S	Masculine singular	Homem amabilíssimo/ADJ-S
	ADJ-S-F	Feminine singular	Mulher belíssima/ADJ-S-F
	ADJ-S-P	Masculine plural	Barcos rapidísimos/ADJ-S-P
	ADJ-S-F-P	Feminine plural	Fragatas imponentísimas/ADJ-S-F-P
ADJECTIVES AS NOUNS	N	General adjectives	<p>Um pecador/N miserável</p> <p><i>versus</i></p> <p>Um homem bom ou mau, justo ou pecador/ADJ....</p>
ADJECTIVES AS ADVERBS	ADV		Fragata rápida/ADJ-F

				<i>versus</i> Rápida/ADV e sorateiramente/ADV  Homem baixo/ADJ <i>versus</i> Falando baixo/ADV, (...)
PARTICIPLES AS ADJECTIVES	VB-AN			Esta ovelha perdida/VB-AN-F  Bemdito/VB-AN e louvado/VB-AN seja Deus Nosso Senhor

TYPE	TAG	VALUE	APPLICATION	EXAMPLES
ADVERBS	ADV	Adverbs of manner		Completamente/ADV Rapidamente/ADV Infelizmente/ADV Deveras/ADV, Depressa/ADV
		Adverbs of place		Acima/ADV, abaixo/ADV, por/P baixo/ADV, em/P cima/ADV, debaixo/ADV, diante/ADV, adiante/ADV, cá/ADV, lá/ADV, aqui/ADV, ali/ADV, aquém/ADV, além/ADV, atrás/ADV, dentro/ADV, fora/ADV (see fora as <a href="#">preposition</a> ),
		Adverbs of time		Agora/ADV, ontem/ADV, hoje/ADV, amanhã/ADV, cedo/ADV, tarde/ADV, jamais/ADV, logo/ADV, antes/ADV, primeiro (=antes)  Antes/ADV d'hontem/P+ADV fui informado.....(see 'antes' as ' <a href="#">comparative adverb</a> ' and <a href="#">more P+ADV</a> )  Primeiro/ADV que/C os vossos achques me viessem (...)
		Others 'acaso' as 'casualmente'		...que quase/ADV interrompemos nós....
EXCAMATIVE/ COMPARATIVE ADVERBS	ADV-R	'mais'/'menos'/ (in all syntactic contexts)  'demais'/'bastante'		por mais/ADV-R força <b>NEW</b> <a href="#">que/C</a> aplicassem, não resolviam...  por mais/ADV-R <b>NEW</b> <a href="#">que/C</a>

			<p>aplicassem força,....</p> <p>O mais/ADV-R são detalhes que...</p> <p>que representam no que avultam mais/ADV-R do que/WPRO custam no que oprimem.</p> <p>Nada procuro mais/ADV-R <del>NEW</del> <a href="#">que/C</a> o amor de Cristo.</p> <p>Por mais/ADV-R <del>NEW</del> <a href="#">que/C</a> eu queira, (...)</p> <p>Pelo menos/ADV-R, isto foi feito.</p> <p>Ao menos/ADV-R isso (...)</p> <p>(...) assim nisto como no demais/ADV-R (...)</p>
		'melhor'/'pior'	....e correu melhor/ADV-R do que eu
		'tão'/'tanto'/'tal'	<p>E fico tão/ADV-R consolado com os favores do Céu, <del>NEW</del> <a href="#">que/C</a> os maiores Impérios da terra regeitara eu.</p> <p>Só sinto ser-lhe tão/ADV-R ingrato e tão pobrezinho para pagar-lhe, <del>NEW</del> <a href="#">que/C</a> nem um suspiro tenho de meu para oferecer-lhe;</p> <p>Quanto mais quer, tanto/ADV-R mais faz para obter</p> <p>Tanto/ADV-R ricos, quanto/WADV pobres...</p> <p>tal/ADV-R qual/WADV sou</p>
		'antes'	Antes/ADV-R tarde do/P+D que/WPRO nunca
SUPERLATIVE ADVERBS	ADV-S		Lentíssimamente
ADVERBS AS PREPOSITIONAL OR ADVERBIAL PHRASES	<p>.../P .../N</p> <p>.../P .../ADJ</p> <p>.../P.../ADV</p>		<p>Com/P certeza/N, , sem/D dúvida/N, por/P acaso/N</p> <p>Por/P conseguinte/ADJ, de/P certo/ADJ ,</p>

	.../ADV.../P		Até/P amanhã/ADV, até/P logo/ADV, <b>NEW</b> no/P+D entanto/ADV ( <a href="#">see clusters</a> )  <b>NEW</b> acerca/ADV de/P ( <a href="#">see clusters</a> )
ADVERBS AS EVENT QUANTIFIERS	Q	'muito'/'pouco'	Frei António trabalhou muito/Q durante ...
<b>NEW</b> NEGATIVE ADVERBS	ADV-NEG	'nunca'	see <a href="#">negation</a>
TYPE	TAG	VALUE/APPLICATION	EXAMPLES
QUANTIFIERS	Q	Quantifiers in general	Tudo/Q será feito.  Alguém/Q disse que...  Outrém/Q  Vossa Mercê sofrerá muito/Q.  muito/Q trabalho a fazer...  pouco/Q trabalho a fazer...  Um pouco/Q de trabalho a fazer  Algun/Q motivo apresenta...
	Q-F		por alguma/Q-F razão...
	Q-P		ambos/Q-P os freis...  todos/Q-P os homens...  alguns/Q-P trabalhos....
	Q-F-P		ambas/Q-F-P as religiosas...  todas/Q-F-P as coisas...  algumas/Q-F-P guerras....
	Q-G		Cada/Q-G um saberá...  Qualquer/Q-G solução virá....
	Q-G-P		Quaiquer/Q-G-P resoluções...
<b>NEW</b> NEGATIVE QUANTIFIERS	Q-NEG  Q-NEG-P, Q-	'nada', 'ninguém', 'nenhum'  ' nenhuns'; 'nenhuma';	se devem vigiar mais, que nenhumas/Q-NEG-F-P outras

	NEG-F, Q-NEG-F-P	'nenhumas'	see <a href="#">negation</a>
TYPE	TAG	VALUE/APPLICATION	EXAMPLES
COORDINATING CONJUNCTIONS	CONJ	Aditive <i>e</i>	Frei Luís e/CONJ Frei João Teria chegado e/CONJ saído sem ninguém perceber....
		Alternatives	Ou....., ou..... Ora....., ora..... Nem....., nem.....
	<b>NEW</b> CONJ-NEG	Aditive <i>nem</i>	Não fez o pedido, nem/CONJ-NEG sabia dizer quem o fizera....  (see <a href="#">negation</a> )
Complementizer	C	'que'	Para dizer a Vossa Mercê que/C esta vida/N entra a ser exemplo de emendas...  Espero que/C Deus me guie....  Foi ontem que/C encontrei as...  Assim/ADV que/C ./, do dinheiro que Vossa Excelência me remeteu não posso já fazer cousa alguma, por estarem as letras nas mãos de quem vendeu a fragata.  (see 'que' as <a href="#">part of complex subordinating conjunctions</a> )
SUBORDINATING CONJUNCTIONS  NB.: In the parsed files, subordinating conjunctions are tagged as <b>C</b> .	CONJS	Causals, conditionals, finals, temporals, etc.	Quando/CONJS fores embora, ...  Embora/CONJS não saibas,....  Se/CONJS soubesses....  Exceto/P se/CONJS...

			<p>Caso/CONJS não saibas.....</p> <p>Conforme/CONJS afirmam,...</p> <p>Como/CONJS afirmam, ....</p> <p>Segundo/CONJS</p> <p>Porque/CONJS</p>
COMPLEX SUBORDINATING CONJUNCTIONS	Various	Adverbs/Prepositions/ Participles/Focus Particle + C	<p>já/ADV que/C</p> <p>para/P que/C</p> <p>visto/VB-AN que/C</p> <p>ainda/ADV que/C</p> <p>depois/P que/C</p> <p>mesmo/FP que/C</p> <p>até/P que/C</p>
QUE	degree clauses		<p>E aplicou tantos/ADJ-R-P esforços que/C impediu o levantar da barreira</p> <p>Fez um tal/ADJ-R esforço que/C não pode....</p> <p>E fico tão/ADV-R consolado com os favores do Céu, que/C os maiores Impérios da terra regeitara eu.</p> <p>por mais/ADV-R força que/C aplicassem, não resolviam...</p> <p>Por mais/ADV-R que/C eu queira, (...)</p>
	comparative clauses		...projetos maiores <b>que/C</b> aqueles....

TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
RELATIVE ELEMENTS	WPRO		<p>O homem que/WPRO veio...</p> <p>O modo pelo qual/WPRO...</p>
	WPRO-P		<p>As casas nas quais/WPRO-P....</p> <p>tantos.....quantos/WPRO-P forem os interessados nas companhias</p>
	WPRO-F-P		<p>não me soube dizer Jerónimo Nunes quantas/WPRO-F-P eram</p>
	WPRO\$		O homem cujo/WPRO\$ caráter....



	WPRO\$-F		...a consciência, cuja/WPRO\$-F glória é aquietar entre as espinhas.
	WPRO\$-P		O homem cujos/WPRO\$-F modos...
	WPRO\$-F-P		...as consciências, cujas/WPRO\$-F-P glórias é aquietar entre as espinhas.
	WADV		O modo como/WADV vejo a Deus é tão sublime que.....  O deserto onde/WADV me encontrava, quando....  Quanto/WADV à compra dos navios,
EXCLAMATIVE/ INTERROGATIVE ELEMENTS	WPRO		Por/P que/WPRO não me deste...?  O/D que/WPRO queres...?
	WPRO-F	quanta	
	WPRO-P	quantos	
	WPRO-F-P	quantas	...porque não me soube dizer Jerónimo Nunes quantas/WPRO-F-P eram....
	WADV		E para saber isto como/WADV é, tende sentido...  Quão/WADV verdadeiro é o sentimento (...)
	WQ		E perguntavas se/WQ não havias de se entregar.
EXCLAMATIVE/ INTERROGATIVE DETERMINERS	WD	Que processo	Qual/WD processo/N procuras?
	WD	Quanto sofrimento	Que/WD processo/N procuras?
	WD	Qual processo	Quanto/WD desamor/N encontrastes, quando.....?
	WD	Que processos	Que/WD processos/N-P...?
	WD-F	Quanta dúvida	terá quanta/WD-F gente quizer
	WD-P	Quantos/quais processos	Quais/WD-P processos/N-P...?  Quantos/WD-P desamores/N-P encontrastes, quando...?

			Os/D-P quais/WD-P livros/N-P....
	WD-F-P	Quantas	Quantas/WD-F-P esperanças/N-P...!
OTHERS	WADV		Quanto/WADV a/P Vossa Mercê, peço a Deus que..... (see <a href="#">prepositions + WADV</a> )
TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
PREPOSITIONS	P	General prepositions:  a, de, com, em, para, até, sobre, desde, exceto, fora (=exceto), salvo, mediante, etc.	
CONTRACTIONS	P+D  (P+D-P; P+D-F; P+D-F-P; P+D-UM; P+D-UM-P; P+D-UM-F; P+D-UM-F-P)	à(s), ao(s), no(s), na(s), do(s), da(s), desse(s), dessa(s), pelo(s), pela(s)	
	P+PRO	dele, dela, etc	
	P+OUTRO  (P+OUTRO-P; P+OUTRO-F; P+OUTRO-F-P)	d'outro (s), doutro(s), doutra, noutra, etc.	quando uma consoante se converte noutra/P+OUTRO-F

	P+Q (P+Q-P; P+Q-F; P+Q-F-P)	n' algum, n' alguns, n' alguma, n' algumas	
	P+WPRO	porque (in interrogatives)	¿ porque/P+WPRO o não fará ?
	P+DEM	disso, disto, nisso, nisto	
	P+ADV	até'gora, antes d'hontem, etc.	
	P+WADV	donde, aonde, adonde	
	P+CL	pela, polo, etc.	trabalhei pela/P+CL servir
	P+NPR	del rey	

ITEM	OLD TAG	<b>NEW</b> NEW TAG
apesar de	P-21 ... P-22  apesar/P-21 de/P-22	<b>P ... P</b>  apesar/P de/P  apesar/P do/P+D; apesar/P dos/P+D- P; apesar/P da/P+D-F; apesar/P das/P+D-F-P; apesar/P disso/P+DEM; apesar/P dele/P+PRO, etc.
acerca de	P-21 ... P-22  acerca/P-21 de/P-22	<b>ADV ... P</b>  acerca/ADV de/P

		acerca/ADV do/P+D; acerca/ADV dos/P+D-P; acerca/ADV da/P+D-F, acerca/ADV das/P+D-F-P; acerca/ADV disso/P+DEM; acerca/ADV dele/P+PRO, etc.
no entanto	CONJ-21 CONJ-22 no/CONJ-21 entanto/CONJ-22	<b>P+D ... ADV</b> no/P+D entanto/ADV
não obstante	CONJ-21 CONJ-22 não/CONJ-21 obstante/CONJ-22	<b>NEG ... ADJ-G</b> não/NEG obstante/ADJ-G

TYPE	TAG	VALUE/APPLICATION	EXAMPLES
OUTRO	OUTRO	Outro <i>(in all contexts)</i>	Foram ao encontro um do outro/OUTRO  O outro/OUTRO homem que partiu...
	OUTRO-P OUTRO-F OUTRO-F-P	Outros Outra Outras <i>(in all contexts)</i>	Acabava por não saber como agradá-los: a uns oferecia amor, a outros/OUTRO-P dedicação.....  Abraçaram-se uns aos outros/OUTRO-P  Uma e outra/OUTRO-F coisa se deve evitar
FOCUS PARTICLES	FP	'só', 'mesmo', 'até', and so on	Só/FP os padres podem.... <b>versus</b> Um só/ADJ livro foi enviado por....  Até/FP Sua Senhoria diria que... <b>versus</b> se pudesse ir até/P a presença de Sua Senhoria, ...
CARDINAL NUMBERS	NUM	General cardinal number expressed by digits or characters	Em 1538/NUM....  Três/NUM toneladas....  100/NUM quilos....  um/D-UM milhão/N

			e/CONJ dois/NUM mil/NUM florins/N-P
	NUM-F	The form 'duas'/2	Duas/NUM-F centenas.... Por 20/NUM homens e 2/NUM-F fragatas....
TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
NEGATIVE PARTICLES	NEG	'não'	Não/NEG se fará.....  dizem/VB-P um/D-UM não/NEG que/WPRO parece/VB-P sim/ADV  não/NEG obstante/ADJ ( <a href="#">see clusters</a> )
	SENAO	'senão'	Não respira senão/SENAO morte.....
OTHER NEGATIVE ITEMS	CONJ-NEG	'nem'  NEW in all contexts	Não fez o pedido, nem/CONJ-NEG sabia dizer quem o fizera....  Nem/CONJ-NEG só de pão vive um ser...
	ADV-NEG	'nunca'	
	Q-NEG	'nada', 'ninguém', 'nenhum'	
	Q-NEG-P, Q-NEG-F, Q-NEG-F-P	' nenhuns'; 'nenhuma'; 'nenhumas'.	

TYPE	TAG	VALUE/APPLICATION	EXAMPLES
INTERJECTIONS	INTJ	General interjections	Oh/INTJ
TYPE	TAG	VALUE/APPLICATION	EXAMPLES
FOREIGN WORDS	FW	General foreign words	repartindo-se/ pro/FW rata/FW  Pax/FW Christi/FW
UNKNOWN WORDS	XX	General unknown words	
TYPE	TAG	VALUE/ APPLICATION	EXAMPLES
FINAL PUCTUATION	.	Period or reticence	./.
		Interrogative mark	?/.

		Exclamation mark	!/.
		Semi-comma	;/.
		Colon	:/.
NON-FINAL PUNCTUATION	,	Comma	,/.
QUOTATION	QT <b>NEW</b>	Quotation mark	"/QT
PARENTHESES	(	Parentheses	((.....))
		Dash	-(
LEXICAL ITEM	TAGS		EXAMPLES
ATÉ	/P /FP		meter a mão até/P o cotovello pelo alheo  até/FP nas arvores há ladroens
COMO	/C /WADV		
MELHOR/PIOR	/ADJ-R-G (with inflectional tag -P, when necessary) /ADV-R (always with no inflectional tags)		
MESMO	/ADJ (with inflexional tags (-F or -P), when necessary) /FP (always with no inflectional tags)		
PORQUE	/C /WADV /P+WPRO		
QUE	/C, /WPRO		
TAL	/D-G (with inflexional tag -P, when necessary) /DEM (always with no inflectional tags) /ADJ-R-G (with inflexional tag -P, when necessary) /ADV-R (always with no inflectional tags)		seus primos, filhos do tal/D-G tio  quem tal/DEM fez, que não faria ?  de tal/ADJ-R-G qualidade , que não he

		possível referillas todas  e mostra-se , que tal/ADV-R deve ser a paz
TANTO	/ <b>ADJ-R</b> (with inflectional tags (-F or -P), when necessary) / <b>ADV-R</b> (always with no inflectional tags)	
LEXICAL ITEM	TAG	
MAIS/MENOS	<b>ADV-R</b>	
NADA	<b>Q-NEG</b>	
NÃO	<b>NEG</b>	
NEM	<b>CONJ-NEG</b>	
NINGUÉM	<b>Q-NEG</b>	
NENHUM	<b>Q-NEG</b>	
nenhuns nenhuma nenhuma	(with inflectional tags –P and/or F, when necessary)	
NUNCA	<b>ADV-NEG</b>	
OUTRO	<b>OUTRO</b>	
outros outra outras	(with inflectional tags –P and/or F, when necessary)	
TÃO	<b>ADV-R</b>	