

Universidade de São Paulo - USP  
Universidade Federal de São Carlos - UFSCar  
Universidade Estadual Paulista - UNESP

## **VERBOS AUXILIARES NO PORTUGUÊS DO BRASIL**

Magali Sanches Duran

Sandra Maria Aluísio

**NILC-TR-10-05**  
**Dezembro 2010**

Série de Relatórios do Núcleo Interinstitucional de Linguística  
Computacional

**NILC - ICMC-USP, Caixa Postal 668, 13560-970 São  
Carlos, SP, Brasil**

# **VERBOS AUXILIARES NO PORTUGUÊS DO BRASIL**

## VERBOS AUXILIARES NO PORTUGUÊS DO BRASIL

Resumo. A anotação de relações semânticas entre um predicado verbal e seus argumentos exige, em primeiro lugar, que se delimite a fronteira entre as partes que serão relacionadas. O predicado verbal pode ser constituído por um único verbo ou por uma sequência de verbos, como “tinha sido informado”, na qual um dos verbos é o núcleo e os demais são auxiliares. Neste relatório descrevemos o estudo sobre essas sequências de verbos, desde sua identificação até seu tratamento nas tarefas de anotação de perguntas em textos e anotação de papéis semânticos. A pesquisa foi realizada por meio de um estudo de corpus e resultou na construção de uma tabela de verbos auxiliares do português do Brasil.

### 1. INTRODUÇÃO

O fenômeno da auxiliação verbal, ou seja, de um verbo auxiliar o outro, há muito é discutido dos linguistas. Benveniste (1965) examinou as relações de auxiliação e observou que um verbo assume funções gramaticais de auxiliar diante da presença de um auxiliado. O fenômeno da auxiliação não está, portanto, somente no verbo auxiliar, mas sim na combinação entre dois verbos, uma forma que auxilia e uma forma que é auxiliada.

Há diferentes formas de auxílio que um verbo pode prestar ao outro e isso varia entre as línguas. No português, temos auxiliares de tempo, modo, aspecto e de diátese (para construção da voz passiva). No inglês, além desses, temos os verbos que auxiliam os outros na construção da forma negativa e interrogativa das orações.

Benveniste chama a atenção para a formação de cadeias de auxiliares, em que um mesmo verbo é auxiliado por mais de um verbo auxiliar. O autor chama a isso *sobreauxiliação* e diz que cada tipo de auxiliação se dá em um nível: a auxiliação de diátese cessa um nível acima da auxiliação de temporalidade e a auxiliação de modalidade cessa um nível acima da auxiliação de temporalidade. (Benveniste, 1965, 2006 p.193.).

*Ele deveria ter sido preso quando tentou deixar o país.*

O mesmo autor aponta que pode haver também uma sobreauxiliação de modalidade (dois auxiliares de modo), exprimindo um alto grau de probabilidade:

(1) *Ele deve poder fazer esse trabalho.*

Os auxiliares de tempo no português são os mais gramaticais, ou seja, não têm mais nenhum conteúdo semântico próprio. Já os auxiliares de modo e aspecto ainda preservam parte de sua semântica, mas têm estabilidade de sentido, ou seja, uma vez usados como auxiliares, geram pouca ou nenhuma ambiguidade de sentido.

O caso mais prototípico de auxiliação no português é o dos verbos *ter* e *haver*, utilizados para formar os tempos compostos.

Exemplos<sup>1</sup>:

(2) *Ele havia comentado isso na última conversa que tivemos.*

(3) *Amanhã ele já terá esquecido tudo o que disse.*

O comportamento dos verbos *ter* e *haver* usados como auxiliares é tomado como modelo pelos linguistas (como Gonçalves, 2002 e Reis, 2010) para discutir se outros verbos podem ou não ser igualmente classificados como auxiliares. As diferenças de comportamento são motivo para que alguns utilizem termos como *verbos semi-auxiliares* e *verbos funcionais* para referenciar os verbos que têm características próximas mas não idênticas às dos verbos *ter* e *haver*.

Para analisar se um verbo pode ser classificado como auxiliar, os linguistas fazem alguns testes. Primeiramente, só são candidatos ao teste aqueles verbos que ocorrem frequentemente seguidos de outro verbo em uma das formas nominais (infinitivo, gerúndio e particípio), pois toda auxiliação implica que apenas o primeiro verbo da sequência seja flexionado. Os três testes mais comuns são:

I. O sujeito do auxiliar tem que ser o mesmo do auxiliado, o que elimina, por exemplo:

(4) *Ele mandou fazer um terno de linho.*

(5) *Ele mandou que fizessem um terno de linho.*

---

<sup>1</sup> Os exemplos que não tiverem fonte de corpus mencionada foram criados apenas para fins de explicação.

Nesse caso, o sujeito de *mandar* é diferente do sujeito de *fazer* e, por isso, é possível transformar a segunda oração em uma relativa, “quebrando” a sequência verbal.

Outro falso positivo dessa categoria é o verbo *fazer* em seu emprego causativo:

(6) *O vento fez voar os chapéus.*

(7) *O vento fez com que os chapéus voassem.*

II. Os modificadores (de negação, de tempo, de lugar etc.) têm que modificar todos os verbos ao mesmo tempo, o que elimina, por exemplo:

(8) *João quer fazer plantão no próximo sábado.*

(9) *Hoje João quer fazer plantão no próximo sábado, mas amanhã talvez mude de ideia.*

(10) *João quer não viajar simplesmente, mas viajar em grande estilo.*

O exemplo IIa mostra o caso do verbo *querer*, que é frequentemente seguido de outro verbo no infinitivo e parece ter uma função de modalizador da ação (e por isso já houve quem o classificasse como verbo auxiliar de modo volitivo). Na primeira oração da sentença (9) há dois modificadores de tempo: *hoje* e *no próximo sábado*, o primeiro modifica o verbo *querer* e o segundo o verbo *fazer*. Já no exemplo (10) o *não* modifica apenas o verbo *viajar* e não o verbo *querer*. Os exemplos (9) e (10) são usados para mostrar que o verbo *querer* admite modificação separada do verbo de sua sequência e, por isso, não pode ser classificado como auxiliar.

Outros falsos positivos dessa categoria são fornecidos por verbos como: *desejar*, *esperar*, *necessitar*, *precisar*, *saber*, *tentar*, *procurar* (no sentido de tentar) etc.

III. A sentença não pode ser parafraseada por uma oração desenvolvida, o que elimina, por exemplo:

(11) *Ele pensou estar no paraíso.*

(12) *Ele pensou que estava no paraíso.*

O exemplo (11) traz um infinitivo que não é auxiliado pelo verbo à esquerda: ele é o complemento objeto direto do verbo à esquerda: um complemento oracional sob forma de oração reduzida de infinitivo. Em (12) mostra-se a paráfrase da oração reduzida por uma oração desenvolvida.

Falsos positivos dessa categoria incluem, entre outros, os verbos: *dizer*, *estimar* (no sentido de fazer estimativa), *julgar*, *jurar*, *prometer*.

A identificação de verbos auxiliares, portanto, deve partir da análise de sequências de verbos. Dada uma sequência, é preciso fazer testes, como os apresentados acima, para decidir se tal sequência forma uma unidade (um sintagma verbal) ou não.

O estudo dos verbos auxiliares costuma estar ligado aos estudos de gramaticalização (Traugott & Heine, 1991, Heine, 1993). Isso porque um verbo, em seu uso auxiliar, “sai” do léxico para “entrar” na gramática. No léxico, o verbo é pleno semanticamente, já na gramática, ele perde sua carga semântica ou tem um sentido fixo para desempenhar determinada *função*.

A gramaticalização é um processo histórico e social, o que significa dizer que há verbos que não eram usados como auxiliares no passado da língua, mas agora o são (v. Travaglia, 2002, 2003, 2004 e 2007; Barroso, 2007, Pinto, 2007 sobre a gramaticalização de verbos no português do Brasil).

O assunto “verbos auxiliares” é complexo e fornece diversas possibilidades de abordagem. Como o propósito do presente estudo é marginal (foi feito para suprir as necessidades de um projeto principal que faz a anotação de estruturas argumentais em dois corpora de português do Brasil), fizemos uma redução teórica do problema e nos concentramos em obter informações essenciais para nossas tarefas.

Contemporânea à nossa pesquisa foi a publicação do trabalho de Baptista et al. (2010), sobre os auxiliares no português europeu, que também tem por objetivo fornecer recursos léxicos para o PLN e trouxe grande contribuição para nossa pesquisa.

Este relatório está dividido em sete seções além desta: na Seção 2 discutimos os auxiliares sob a ótica do processamento de línguas naturais (PLN), na Seção 3 discutimos o tratamento dos auxiliares pelo parser Palavras (Bick, 2000), que é largamente utilizado em PLN. Discorremos na Seção 4 sobre o tratamento dos auxiliares na anotação de papéis semânticos e, na Seção 5, sobre a anotação de perguntas em texto. Na Seção 6 relatamos o processo de extração dos candidatos no corpus do NILC utilizando a ferramenta AC/DC. Na Seção 7 expomos as contribuições de Baptista et al. (2010) para nosso trabalho e discutimos as diferenças entre os auxiliares nas variantes brasileira e europeia do português. Por fim, na Seção 8 tecemos comentários finais e apontamos possibilidades de trabalhos futuros.

## 2. O FENÔMENO DA AUXILIAÇÃO PARA FINS DE PLN

Para fins de processamento automático de língua, não é possível informar simplesmente que os verbos *ter* e *haver*, por exemplo, são auxiliares, pois esses verbos admitem uso como verbos plenos (*ter*, no sentido de *possuir* e *haver*, no sentido de *existir*, por exemplo). A informação que permite identificar o uso auxiliar desses verbos é a forma do verbo auxiliado. O verbo auxiliado é sempre o verbo à direita do auxiliar e ele sempre assume uma forma nominal exigida pelo verbo auxiliar (infinitivo ou particípio ou gerúndio). No caso de *ter* e *haver*, é possível dizer que estão funcionando como auxiliares sempre que:

- a) Tiverem um verbo à direita E
- b) O verbo à direita estiver na forma do particípio passado.

Além disso, existem sequências mais complexas, formadas por mais de um auxiliar:

(13) *Ele devia ter dito isso antes.*

No exemplo, o verbo *dever* é um auxiliar de modo, o verbo *ter* é auxiliar de tempo e o verbo *dizer* é o auxiliado.

Então, em uma sequência de auxiliares, apenas o primeiro (à esquerda) é flexionado (verbo na forma finita) e os demais assumem a forma nominal exigida pelo verbo auxiliar imediatamente à esquerda.

No caso do exemplo (13), o verbo *dever* impõe a forma do infinitivo ao verbo *ter* e o verbo *ter* impõe a forma do particípio ao verbo *dizer*.

O verbo principal é sempre o último verbo de uma sequência formada por

(VFIN) + VINF + VINF + VINF

Sendo VFIN a forma finita de um verbo e VINF uma das formas nominais ou infinitas (infinitivo, particípio ou gerúndio). Note-se que o primeiro auxiliar pode ou não estar na forma finita, pois há ocorrências como:

(14) *Tendo chegado tarde, ele não ouviu as primeiras palestras.*

(15) *Ele pretendia ter dito isso antes.*

No exemplo (5), embora a sequência de verbos seja VFIN + VINF + VINF, ela não forma um sintagma verbal, pois o verbo *pretender* nunca é classificado como auxiliar (ele admite que apenas o verbo que o segue seja modificado: *Ele pretendia não ter dito isso antes*). A sequência *ter dito* é um VP que funciona como complemento objeto direto de *pretendia*, que é outro VP.

Para identificar as sequências de auxiliação é preciso, além das regras acima, um recurso léxico que informe quais os verbos que podem atuar como auxiliares. Além disso, dado que um verbo nunca é exclusivamente usado como auxiliar, é preciso que esse recurso léxico informe também qual a forma infinita cada verbo auxiliar impõe ao verbo que está à sua direita, pois um verbo pode ser seguido de um verbo em forma infinita que não seja aquela que ele exige para atuar como auxiliar. Por exemplo, o verbo *ser* é auxiliar de diátese (forma a voz passiva) e exige que o auxiliado esteja no particípio passado (exemplo 16). Se o verbo *ser* for seguido de infinitivo (exemplos 17 e 18), ele não deverá ser interpretado como auxiliar, ou seja, ele não formará um VP com o verbo no infinitivo que o segue:

(16) *Ele foi perseguido pelos policiais.*

(17) *O importante é fazer isso sem demora.*

(18) *O que ele deseja é sair sem ser notado.*

Há, portanto, duas informações necessárias em um recurso léxico de apoio à identificação de sintagmas verbais formados por processo de auxiliação:

VERBO	FORMA NOMINAL DO VERBO À DIREITA
-------	-------------------------------------

Isso leva a uma tabela do tipo:

VERBO	FORMA NOMINAL DO VERBO À DIREITA
ter	Particípio
haver	Particípio
ir	Infinitivo
dever	Infinitivo
poder	Infinitivo

Contudo, os verbos auxiliares possuem uma outra característica: cada um presta um tipo de auxiliação: construção do tempo, do modo, do aspecto e da voz passiva. Os auxiliares de tempo auxiliam a localizar temporalmente a ação do verbo principal. Os auxiliares de modo ajudam a dizer se a ação do verbo principal pertence ao universo da possibilidade, da obrigatoriedade, da proibição etc. Os auxiliares de aspecto auxiliam a dizer se a ação do verbo principal está começando (inceptiva), terminando (terminativa) ou se ainda não se realizou (imperfectiva); se tem uma duração (durativa) ou se repete ao longo do tempo (frequentativa), etc.

Benveniste, que não reconhecia ainda os auxiliares de aspecto, observou no francês a ordem de ocorrência dos tipos de auxiliares em sequências de sobreauxiliação. Segundo ele, os auxiliares de MODO precedem os auxiliares de TEMPO que, por sua vez, precedem os auxiliares de DIÁTESE (usados para fazer a voz passiva). No português parece que a mesma ordem é observada:

(19) *Ele pode ter sido feito há anos..*

(20)\**Ele ter pode sido feito...*

(21)\**Ele sido pode ter feito...*

Observação: o símbolo \* marca as sentenças agramaticais.

Se incluirmos um auxiliar de aspecto na sequência, percebemos que ele se coloca entre o auxiliar de tempo e o auxiliar de diátese:

(22) *Ele pode ter começado a ser feito há anos.*

A sequência natural dos verbos auxiliares, portanto, seria:

MODO-TEMPO-ASPECTO-DIÁTESE (MTAD).

Encontramos, contudo, exemplos de sequências que contrariam essa ordem:

*par=7889: Queria ter podido falar isso para ele naquele dia (TM)*

*par=120897: D. Fernanda leu-a; era do diretor da casa de saúde; noticiava que Rubião, desde três dias, desaparecera, não tendo podido ser encontrado por mais esforços que houvessem empregado a polícia e ele .(TMD)*

*par=Especial-94a-nd-1: Isso significa que, pelo menos naquele Estado, a promessa de palanque não vai poder ser cumprida .(TMD)*

Esses exemplos indicam que a ordem de encadeamento dos verbos auxiliares mereceria uma pesquisa específica.

A informação do tipo de auxílio prestado pode ser usada para anotar os auxiliares e, posteriormente, observar as sequências de tipos. Essa informação é muito relevante para fins de geração de língua e, por isso, nossa tabela recebeu uma terceira coluna, que é o tipo de auxiliar:

VERBO	FORMA NOMINAL DO VERBO À DIREITA	TIPO
ter	Particípio	TEMPO
haver	Particípio	TEMPO

ir	Infinitivo	TEMPO
dever	Infinitivo	MODO
poder	Infinitivo	MODO
deixou de	Infinitivo	ASPECTO
ser	Particípio	DIÁTESE

### 3. VERBOS AUXILIARES NO PARSER PALAVRAS

Uma vez que, em nossas tarefas de anotação, estamos usando um corpus anotado pelo parser Palavras (Bick, 2000), herdamos todos os verbos auxiliares já reconhecidos por ele. No entanto, o Palavras não reconhece todos os auxiliares que identificamos ao longo de nosso estudo. Isso significa que os resultados deste estudo poderão ser incorporados ao Palavras e a outros parsers.

Na análise do parser Palavras, o conjunto de um ou mais auxiliares seguidos de um verbo pleno recebe uma etiqueta de VP (Verbal Phrase) ou, seja, sintagma verbal.

Se o Palavras reconhecesse todos os auxiliares, seria simples identificar o verbo principal (o verbo que deve evocar os argumentos, o centro de uma estrutura argumental): seria preciso apenas selecionar o último verbo à direita do VP.

Encontramos também casos em que o Palavras identifica erroneamente VPs. Parece que o parser agrupa sequências de verbos iniciadas por um verbo marcado como auxiliar, como o verbo *ser*, no exemplo abaixo:

*(zh016.s15) Se o sucesso que estamos vendo continuar , será possível manter um nível igual de segurança com menos soldados americanos ∩ disse o presidente .*

Nesse caso, a sequência “estamos vendo continuar” foi analisada como VP pelo Palavras. No entanto, o verbo *ver* nunca é auxiliar e a análise correta seria:

*Se [o sucesso [que estamos vendo]] continuar.*

Portanto, apenas *estamos vendo* é um VP. O verbo *continuar* não pertence à oração encaixada *que estamos vendo*, mas sim à oração *se o sucesso [...] continuar*.

Fizemos uma busca no corpus NILC, disponível para consulta por meio da ferramenta AC/DC (<http://www.linguateca.pt/ACDC/>) a fim de descobrir quais os verbos que foram anotados como auxiliares pelo Palavras. A expressão de busca utilizado foi: **[pos="V" & func="FAUX"]**. Solicitamos a distribuição de lemas e o resultado é o seguinte:

### Distribuição

Houve **23** valores diferentes de **lema**.

ser	39211
poder	24713
ir	16648
ter	14901
estar	14134
dever	11672
haver	4130
vir	3702
começar	2282
acabar	1864
passar	1342
continuar	1203
parecer	1069
chegar	999
deixar	909
costumar	833
voltar	609
ficar	297
parar	203
andar	158
tornar	28
seguir	26
soer	5

Figura 1. Resultado da busca de verbos anotados como auxiliares

Como pode ser verificado, o verbo *ver* não está entre os verbos anotados como auxiliares pelo Palavras e, portanto, o que deve ter gerado sua anotação dentro de um VP na sentença (*zh016.s15*) deve ter sido alguma regra do parser que desconhecemos.

Analisando a lista de auxiliares anotados pelo Palavras e comparando-a com a nossa tabela de auxiliares (Anexo 1), podemos verificar que nossa tabela contém mais

verbos que a lista do Palavras e que todos os verbos auxiliares do Palavras estão contidos em nossa tabela, com exceção do verbo “soer”. Esse verbo aparece na expressão “como sói acontecer” (que significa “como costuma acontecer”), presente nas cinco ocorrências do verbo “soer” computadas no resultado apresentado na Figura 1 (o verbo “soer” tornou-se obsoleto no português).

#### 4. VERBOS AUXILIARES NA ANOTAÇÃO DE PAPÉIS SEMÂNTICOS

A tarefa de anotar um corpus com perguntas ou papéis semânticos (vide Relatório Técnico NILC 03-10) compreende três sub-tarefas:

- a) Identificar o predicado;
- b) Identificar os argumentos, delimitando-os;
- c) Atribuir rótulo mais adequado para unir (a) a (b).

A ordem dessas sub-tarefas deve ser essa, pois o resultado de uma é insumo para a próxima.

A identificação dos verbos auxiliares é uma das dificuldades da primeira tarefa: como reconhecer automaticamente um predicado verbal constituído por um verbo pleno e um ou mais verbos que funcionam como auxiliares?

Segundo Fillmore (1968), na estrutura profunda da língua estão as proposições e em “camadas” mais superficiais encontram-se os modificadores das proposições (a negação, o tempo, o lugar, o modo, o aspecto etc.). Sob essa ótica, os verbos auxiliares não fazem parte das proposições: são modificadores.

Por esse motivo, nossa ideia inicial de selecionar os VPs como evocadores de argumentos não foi mantida. Para fins de anotação semântica, os argumentos pertencem ao verbo principal e não ao VP como um todo.

Sendo assim, é preciso decompor os VPs já reconhecidos pelo parser a fim de fazer a anotação semântica: os auxiliares são anotados como argumentos modificadores

(argumentos externos e posteriores à proposição) e os verbos principais são os evocadores dos verdadeiros argumentos da proposição (argumentos internos).

No Propbank (Palmer et al., 2005), modelo utilizado para nossa anotação, duas etiquetas de argumentos modificadores foram criadas para anotar os verbos auxiliares: *argm-aux*, para os auxiliares de tempo, negação/interrogação e voz passiva (*do, have e be*) e *argm-mod*, para os auxiliares modais (*will, may, can, must, shall, might, should, could, would, going to, have to e used to*). Os verbos aspectuais não são anotados separadamente no Propbank. Sua ocorrência recebe uma anotação de sentido específica e, embora não seja confundido com um verbo pleno, não é anotado como modificador de um verbo pleno.

No nosso projeto Propbank-Br, identificamos automaticamente os verbos auxiliares, de modo que nenhum deles gerasse instância de anotação. A anotação desses verbos (Figura 2) é feita também automaticamente, com apoio da tabela de auxiliares, utilizando etiquetas específicas para auxiliares de tempo, modo, aspecto e voz passiva: *argm-aux-tmp*, *argm-aux-mod*, *argm-aux-asp*, *argm-aux-pas*.

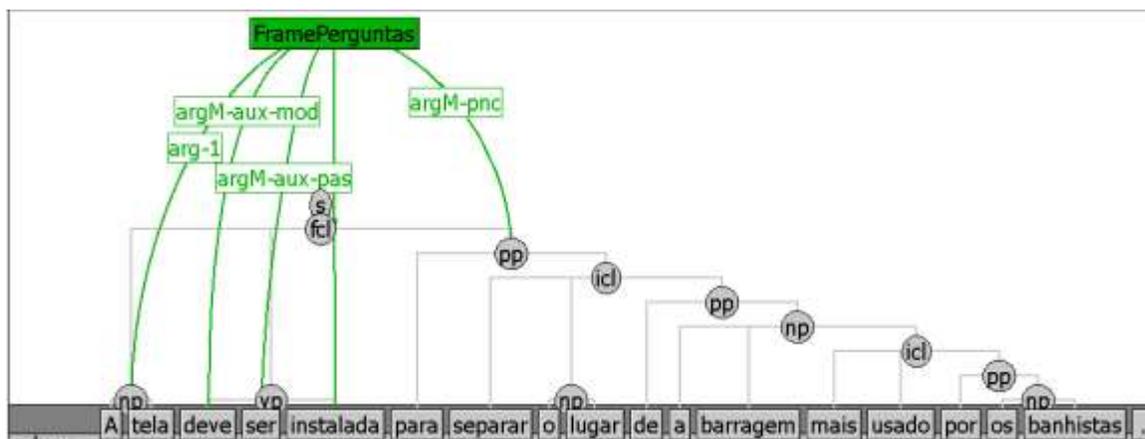


Figura 2: Anotação de verbos auxiliares na tarefa de anotação de papéis semânticos

## 5. VERBOS AUXILIARES NA ANOTAÇÃO DE RÓTULOS DE PERGUNTAS

O tratamento dos verbos auxiliares na anotação de rótulos de perguntas em textos é diferente do tratamento dado a eles na anotação de papéis semânticos.

Na anotação de perguntas, o “centro” da estrutura que liga o predicado aos argumentos com rótulos de perguntas é um VP.

É preciso manter o VP integral nesse tipo de anotação (Figura 3), pois todos os auxiliares serão utilizados, juntamente com o verbo principal, para gerar as perguntas em língua natural.

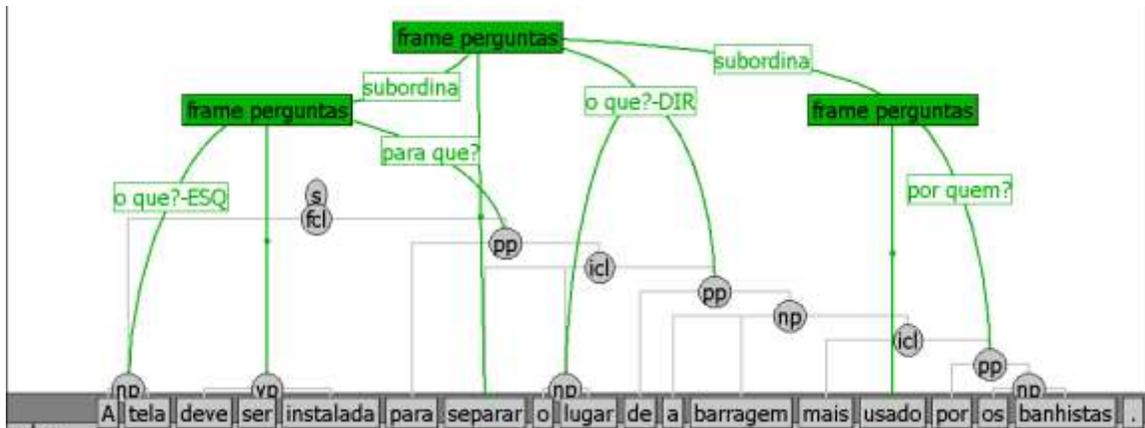


Figura 3: Anotação de verbos auxiliares na tarefa de anotação de perguntas em texto.

A partir da anotação mostrada na Figura 3, serão geradas várias perguntas, pois a sentença tem três verbos evocando perguntas (instalar, separar e usar). O verbo *instalar* nessa sentença tem dois auxiliares: um de modo (*dever*) e um de diátese (*ser*), pois está na voz passiva. As perguntas geradas para o verbo *instalar* são:

1. O que deve ser instalado? Resposta: A tela
2. Deve ser instalada para quê? Resposta: Para separar o lugar da barragem mais usado pelos banhistas.

As formas nominais do verbo, infinitivo, particípio e gerúndio, são chamadas “nominais” porque têm valor de nome: o infinitivo quase sempre tem valor de substantivo, o particípio, de adjetivo e o gerúndio, de advérbio. Por isso, é preciso distinguir se as formas nominais fazem parte do núcleo do predicado e, portanto, serão utilizadas para construir as perguntas, ou se fazem parte dos argumentos do verbo e, portanto, serão utilizadas para construir as respostas. Por exemplo:

No entanto, quando há elipse do verbo auxiliar em uma das orações da sentença, a anotação fica prejudicada, como no exemplo a seguir:

fsp002.s18: Um de os telescópios está pronto e funcionando em o Havaí , EUA .  
Nesse exemplo ocorre uma elipse do verbo *estar* na segunda oração: “está pronto e [está] funcionando.” Enquanto “está pronto” é um predicado nominal, formado pelo verbo de ligação mais adjetivo, “está funcionando” é uma locução verbal de tempo contínuo e teria sido classificado como VP pelo parser se o verbo não estivesse elíptico.

- a. Se anotamos o segmento “está pronto e funcionando”, podemos perguntar: “O que está pronto e funcionando?” e a resposta: “Um dos telescópios”, assim como podemos perguntar “Está pronto e funcionando onde?” e a resposta: “No Havaí, EUA”.
- b. Se anotamos apenas “está”, a pergunta seria: “Está como?” e a resposta: “Pronto e funcionando no Havaí, EUA”

Optamos por anotar como descrito na opção a, mas o ideal seria dar um tratamento às elipses, solucionando-as em uma fase pré-anotação, pois aí teríamos duas orações completas:

1. Um dos telescópios está pronto e
2. [um dos telescópios] [está] funcionando no Havaí, EUA.

Pelo que pudemos observar na anotação do Propbank, o PennTreeBank tem esse tipo de resolução, bem como a resolução de anáforas e correferências.

## 6. IDENTIFICAÇÃO DE VERBOS AUXILIARES POR MEIO DE PESQUISA EM CORPUS

Verbos auxiliares não são exclusivamente auxiliares, eles “estão” auxiliares. Uma das formas mais básicas para se encontrar verbos auxiliares é procurar ocorrências de dois ou mais verbos em sequência. Além disso, cada verbo auxiliar exige que o verbo auxiliado venha à direita e assuma uma das três formas nominais do verbo: infinitivo, gerúndio ou particípio passado.

Esse foi o primeiro argumento de busca que utilizamos em nossa pesquisa, usando o corpus do NILC e a ferramenta AC/DC disponíveis na Linguateca<sup>2</sup>. Colocamos como critério de busca:

VFIN @ VINF

VFIN @ VPCP

VFIN @ VGER

Onde VFIN é um verbo na forma finita, VINF, um verbo no infinitivo, VPCP, um verbo no particípio passado e VGER, um verbo no gerúndio.

Ao analisarmos o resultado desse tipo de busca em corpus, percebemos que esse critério não produz um resultado totalmente “limpo”, pois muitos verbos seguidos de infinitivo, gerúndio e particípio passado não são ou não estão exercendo um papel de auxiliar. Os motivos para isso são:

a. Orações subordinadas reduzidas

*par=9716: Maria disse [ter sido acusada] de furto por Vera .*

*par=7889: Queria [ter podido falar] isso para ele naquele dia .*

*Cotidiano-94b-soc-2: Em depoimento à polícia, ela [havia dito][ter mandado] [matar] seus pais . (=havia dito que tinha mandado que matassem)*

b. Adjetivos formados por participípios

*par=Mundo-94a-pol-2: Mas em Washington é proibido andar armado.*

c. Adjuntos adverbiais formados por gerúndio ou particípio

*=290: Os reflexos condicionados podem ser obtidos associando-se um novo estímulo sensitivo a um estímulo natural que provoque um reflexo natural.*

*par=6355: Os íons sódio também podem ser encontrados associados ao íon cloro e as proteínas .*

---

<sup>2</sup> <http://www.linguateca.pt/cetenfolha/>

Não é possível distinguir automaticamente os verdadeiros VPs de ocorrências como as exemplificadas nos itens a. b. e c. Para melhorar a precisão do reconhecimento de locuções verbais, foi preciso, portanto, construir uma tabela de verbos auxiliares. Essa tabela permite que os verbos auxiliares sejam reconhecidos e integrados aos verbos auxiliados, formando um só VP.

Utilizando diversos argumentos de pesquisa, encontramos em nossas buscas VPs formados por até três verbos auxiliares:

*par=10185: Já esse César Maia vai acabar sendo eleito governador.*

No entanto, utilizando-se um corpus maior, acreditamos que possam ser encontrados casos de quatro e até cinco auxiliares, como:

*(9) Ele podia ter acabado de ser examinado quando você chegou.*

*(10) Ele devia ter podido começar a ser examinado antes das 8h. da manhã.*

## 7. AUXILIARES NAS VARIANTES EUROPÉIA E BRASILEIRA DO PORTUGUÊS

Ao lermos o trabalho de Baptista et al. (2010), percebemos imediatamente a grande sobreposição de nossas pesquisas. A diferença fica por conta dos métodos utilizados. A pesquisa dos autores foi realizada em um corpus muito maior e utilizou técnicas automáticas de extração de sequências de verbos. A análise dos candidatos foi manual, como no nosso caso.

A tabela produzida por Baptista et al. não inclui o que chamamos de “auxiliar de diátese”. Os autores analisam a voz passiva como um verbo de cópula seguido de um adjetivo formado pelo verbo no particípio, que tem a função de predicativo.

Encontramos, além dessa, diferenças entre as tabelas com relação a:

1. Verbos auxiliares
2. Preposição que acompanha o verbo auxiliar
3. Forma nominal exigida pelo auxiliar.

Quanto aos verbos auxiliares identificados, alguns são observados nas duas variantes, outros são observados apenas em uma das variantes. Por exemplo, na variante brasileira temos alguns os verbos aspectuais não identificados na variante europeia:

1. Estar para + infinitivo (*Ele está para completar 10 anos*) e
2. Sair + Gerúndio (*Ele saiu comprando tudo*).
3. Dar de + Infinitivo (*Ele deu de sair toda noite desde que se separou*)
4. Dar para + Infinitivo (*Ele deu para mentir nos últimos tempos*)

Observamos também que a tabela dos autores contém uma coluna a mais, com a classificação da função dos tipos de auxiliares. Fizemos uma versão de nossa tabela com essa nova coluna também, mas não estamos utilizando-a por enquanto em nossas tarefas.

As diferenças entre as tabelas podem ser verificadas em detalhe no Anexo 2, tabela que combina os auxiliares nas duas variantes do português.

## 8. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Para a finalidade a que nos propusemos, a pesquisa sobre auxiliares foi plenamente satisfatória e contribuiu para a automação de alguns passos de nosso processo de anotação de papéis semânticos e de perguntas em textos.

Contudo, a pesquisa foi feita em um corpus relativamente pequeno e poderia ser repetida de forma mais sofisticada em um corpus maior.

O AC/DC é uma ferramenta acessível para pesquisas linguísticas, mas nesse caso muitas buscas tiveram que ser feitas para se chegar aos resultados. O uso de programação teria abreviado o tempo das buscas e permitido o uso de um corpus maior, não disponível para consulta pelo AC/DC, como, por exemplo, o PLN-Br.

Não foi possível também, dadas as limitações da ferramenta AC/DC, pesquisar as sequências de combinação de tipos de auxiliares, informação muito relevante para a geração de língua.

Por esses motivos, acreditamos que os verbos auxiliares merecem uma pesquisa em maior escala, em conjunto com profissionais da computação.

## REFERÊNCIAS BIBLIOGRÁFICAS

BAPTISTA, J.; MAMEDE, N.; GOMES, F. (2010). Auxiliary verbs and verbal chains in European Portuguese. In: Pardo, T.A.S. et al. (Eds.): Computational Processing of the Portuguese Language. *Proceedings of PROPOR 2010*, LNAI 6001, pp. 110–119, 2010.

BARROSO, Paulo Henrique de Oliveira. O verbo *buscar* em processo de gramaticalização – sua recategorização. *Estudos Linguísticos XXXVI(2)*, maio-agosto, 2007. p. 248-257 Disponível em: <http://www.gel.org.br/estudoslinguisticos/edicoesanteriores/4publica-estudos-2007/sistema06/55.PDF>. Acesso em: 26/11/09.

BENVENISTE, E. (2006) Estrutura das relações de auxiliaridade. In: *Problemas de Lingüística Geral II*, cap. 13. p. 181-219. Campinas: Pontes Editores.

BICK, E. (2000). *The Parsing System Palavras Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Aarhus, Denmark, Aarhus University Press.

FILLMORE, C. (1968). The Case for Case. In Bach and Harms (Ed.): *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston, 1-88.

GONÇALVES, Anabela (2002). Verbos auxiliares e verbos de reestruturação do Português Europeu. In Duarte, I. M., J. Barbosa, S. Matos & T. Husgen (orgs.). *Actas do Encontro comemorativo dos 25 anos do Centro de Linguística da Universidade do Porto*. Porto: Centro de Linguística da Universidade do Porto.

HEINE, B. (1993). *Auxiliaries – Cognitive forces and grammaticalization*. Oxford University Press.

PALMER, M. et alli (2005) The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31:1., pp. 71-105, March, 2005.

PINTO, P. J. G. (2007). A (poli)gramaticalização do verbo “deixar”. *Domínios da Linguagem*.

REIS, Fernanda Elena de Barros. auxílio e posição do clítico pronominal no português brasileiro. *Anais do Seta*. Número 4, 2010. Campinas, Instituto de Estudos da Linguagem da UNICAMP. disponível em: <http://www.iel.unicamp.br/revista/index.php/seta/article/viewFile/930/674>. Acesso em: 11/02/2011.

TRAUGOTT, E.C.; HEINE B. (1991). *Approaches to grammaticalization*. Amsterdam: John Benjamins.

TRAVAGLIA, L. C. (2002). *Gramaticalização de verbos – Relatório de pesquisa*. Rio de Janeiro: Faculdade de Letras / UFRJ, Relatório de Pós-Doutorado em Linguística.

TRAVAGLIA, L. C. (2003). A gramaticalização de verbos. In HENRIQUES, Claudio Cezar (org.). *Linguagem, conhecimento e aplicação – Estudos de língua e linguística*. Rio de Janeiro: Editora Europa, p.306-321.

TRAVAGLIA, L. C. (2004). A poligramaticalização do verbo *acabar*. *Letras & Letras*, Uberlândia, 2004.

TRAVAGLIA, L.C. (2007) A gramaticalização dos verbos *passar* e *deixar*. *Revista da Abralin*, vol. 6 n. 1 Disponível em: <http://www.abralin.org/revista/RV6N1/01-Luiz-Carlos-Travaglia.pdf> Acesso em: 26/11/09.

# ANEXO 1

## TABELA DE VERBOS AUXILIARES

VERBO	PREP	FORMA DO AUXILIADO	TIPO DE AUXÍLIO	SEMÂNTICA
acabar		gerúndio	aspectual	terminativo
acabar	de	infinitivo	aspectual	terminativo
acabar	por	infinitivo	aspectual	terminativo
andar		gerúndio	aspectual	durativo
andar	a	infinitivo	aspectual	frequentativo
cessar	de	infinitivo	aspectual	terminativo
chegar	a	infinitivo	aspectual	incoativo
colocar-se	a	infinitivo	aspectual	incoativo
começar	a	infinitivo	aspectual	incoativo
começar	por	infinitivo	aspectual	incoativo
continuar		gerúndio	aspectual	durativo
continuar	a	infinitivo	aspectual	durativo
continuar	sem	infinitivo	aspectual	frequentativo negativo
correr	a	infinitivo	aspectual	incoativo
costumar		infinitivo	aspectual	frequentativo
dar	de	infinitivo	aspectual	incoativo
dar	para	infinitivo	aspectual	incoativo
deitar	a	infinitivo	aspectual	incoativo
deixar	de	infinitivo	aspectual	terminativo
deixar (de x para y)	de	infinitivo	aspectual	negativo
desatar	a	infinitivo	aspectual	incoativo
dever		infinitivo	modal	deôntico/epistêmico
disparar	a	infinitivo	aspectual	incoativo
estar	para	infinitivo	aspectual	ação iminente
estar		gerúndio	aspectual	durativo
estar	sem	infinitivo	aspectual	imperfectivo
estar	por	infinitivo	aspectual	imperfectivo/passiva
estar		particípio	diátese	
estar a ponto de		infinitivo	aspectual	ação iminente
estar em vias de		infinitivo	aspectual	ação iminente
estar prestes a		infinitivo	aspectual	ação iminente
ficar		gerúndio	aspectual	durativo
ficar	sem	infinitivo	aspectual	terminativo
ficar	de	infinitivo	modal	deôntico
haver	de	infinitivo	modal	deôntico
haver	que	infinitivo	modal	deôntico/impessoal
haver		particípio	temporal	passado
ir		gerúndio	aspectual	durativo
ir		infinitivo	temporal	futuro
parar	de	infinitivo	aspectual	terminativo
passar	a	infinitivo	aspectual	incoativo

permanecer		infinitivo	aspectual	durativo
poder		infinitivo	modal	deôntico/epistêmico
pôr-se	a	infinitivo	aspectual	incoativo
precisar		infinitivo	modal	deôntico
recomeçar	a	infinitivo	aspectual	incoativo/iterativo
sair		gerúndio	aspectual	incoativo
seguir		gerúndio	aspectual	durativo
ser		particípio	diátese	
ter	de	infinitivo	modal	deôntico
ter	que	infinitivo	modal	deôntico
ter		particípio	temporal	passado
terminar		gerúndio	aspectual	terminativo
tocar	a	infinitivo	aspectual	incoativo
tornar	a	infinitivo	aspectual	incoativo/iterativo
vir		gerúndio	aspectual	durativo
vir	a	infinitivo	aspectual	terminativo
viver		gerúndio	aspectual	frequentativo
viver	a	infinitivo	aspectual	frequentativo
voltar	a	infinitivo	aspectual	incoativo/iterativo

## ANEXO 2

### TABELA COMPARATIVA DE VERBOS AUXILIARES PORTUGUÊS EUROPEU (PB) VERSUS PORTUGUÊS BRASILEIRO (PB)

Variante Européia	Variante Brasileira	Auxiliar	Part .	AUXILIADO	TIPO DE AUXÍLIO	Função
PE	PB	acabar		GER	VASP	terminativo
PE		acabar	a	INF	VASP	terminativo
PE	PB	acabar	de	INF	VASP	terminativo
PE	PB	acabar	por	INF	VASP	terminativo
PE	PB	andar		GER	VASP	durativo
PE		andar	a	INF	VASP	durativo
PE	PB	cessar	de	INF	VASP	terminativo
PE	PB	chegar	a	INF	VASP	terminativo
PE	PB	começar	a	INF	VASP	incoativo
PE	PB	começar	por	INF	VASP	incoativo
PE	PB	continuar		GER	VASP	durativo
PE	PB	continuar	a	INF	VASP	durativo
	PB	continuar	sem	INF	VASP	negativo
PE		correr	a	INF	VASP	incoativo
PE	PB	costumar		INF	VASP	frequentativo
PE		costumar	de	INF	VASP	frequentativo
	PB	dar	de	INF	VASP	incoativo
	PB	dar	par a	INF	VASP	incoativo
PE	PB	deitar	a	INF	VASP	incoativo
PE		deitar-se	a	INF	VASP	incoativo
	PB	deixar	de	INF	VASP	negativo
PE	PB	deixar	de	INF	VASP	terminativo
PE	PB	desatar	a	INF	VASP	incoativo
PE	PB	dever		INF	VMOD	deôntico/epistémico
PE	PB	dever	de	INF	VMOD	deôntico/epistémico
	PB	estar		PART	VDIA	passiva de estado
PE	PB	estar		GER	VASP	durativo
PE		estar	a	INF	VASP	durativo
PE	PB	estar	par a	INF	VASP	imperfectivo
PE	PB	estar	por	INF	VASP	imperfectivo/passiva
	PB	estar	par a	INF	VASP	incoativo
	PB	estar	sem	INF	VASP	negativo
PE	PB	estar a ponto de		INF	VASP	incoativo
PE	PB	estar prestes a		INF	VASP	incoativo
PE	PB	ficar		GER	VASP	durativo
PE		ficar	a	INF	VASP	durativo
	PB	ficar		GER	VASP	durativo
	PB	ficar	sem	INF	VASP	negativo
PE	PB	ficar	de	INF	VMOD	obrigação
PE	PB	haver	de	INF	VMOD	deôntico
PE		haver	que	INF	VMOD	deôntico/impessoal
PE	PB	haver		PART	VTEMP	passado

PE		hesitar		INF	VMOD	incoativo
PE	PB	hesitar	em	INF	VMOD	terminativo
PE	PB	ir		GER	VASP	durativo
PE	PB	ir		INF	VTEMP	futuro
PE		ir	a	INF	VTEMP	futuro
	PB	ousar		INF		incoativo
PE	PB	parar	de	INF	VASP	terminativo
PE	PB	passar	a	INF	VASP	incoativo
	PB	permanecer		INF		durativo
PE	PB	poder		INF	VMOD	deôntico/epistémico
PE		pôr-se		GER	VASP	incoativo
PE	PB	pôr-se	a	INF	VASP	incoativo
PE	PB	recomeçar	a	INF	VASP	incoativo/iterativo
	PB	sair		GER	VASP	incoativo
	PB	seguir		GER	VASP	durativo
	PB	ser		PART	VDIA	passiva de ação
PE	??	ser	de	INF	VMOD	imperfectivo/deôntico
PE	PB	ter	de	INF	VMOD	obrigação
PE	PB	ter	que	INF	VMOD	obrigação
PE	PB	ter		PART	VTEMP	passado
	PB	terminar		GER	VASP	terminativo
	PB	terminar	por	INF	VASP	terminativo
PE	PB	tornar	a	INF	VASP	incoativo/iterativo
PE	PB	vir		GER	VASP	durativo
PE	PB	vir	a	INF	VASP	terminativo
PE		vir	de	INF	VASP	terminativo
	PB	vir		INF	VASP	terminativo
	PB	viver		GER	VASP	frequentativo
PE	PB	voltar	a	INF	VASP	incoativo/iterativo