

# Dependência: o conceito e as gramáticas

## Dependence: concept and grammars

Isaac Souza de Miranda Jr.   
Universidade Federal de São Carlos

Oto Araújo Vale   
Universidade Federal de São Carlos

### Resumo

As Gramáticas de Dependência (GD) têm ganhado destaque nas últimas décadas, sobretudo em virtude de sua ampla aplicação no Processamento de Linguagem Natural (PLN). Diversas tarefas de análise sintática automática adotam GD como modelo formal, dada sua eficiência na representação binária das estruturas e na desambiguação sintática (Nivre, 2005; Kübler et al., 2009). Projetos como o Universal Dependencies (De Marneffe et al., 2021) consolidaram as GD como paradigma dominante em sistemas multilíngues de anotação, e trabalhos recentes como (Lopes et al., 2024; Souza et al., 2024; Duran & Pardo, 2024) demonstram sua vitalidade nas pesquisas atuais. Diante disso, este artigo tem como objetivo apresentar os fundamentos formais centrais das GD — conectividade, direcionalidade, hierarquia e rotulação sintática —, discutindo definições estruturais da dependência, os princípios teóricos compartilhados entre diferentes modelos e, por fim, demonstrando sua aplicação por meio de uma versão simplificada da Gramática de Hays e Gaifman (Hays, 1964; Gaifman, 1965). O escopo do trabalho é restrito à caracterização teórica e abstrata das propriedades essenciais das GD, com vistas a fornecer uma base introdutória, porém sistemática, para pesquisadores interessados na abordagem.

### Palavras chave

syntax; gramática de dependência

### Abstract

Dependency Grammars (DGs) have gained prominence in recent decades, especially due to their widespread application in Natural Language Processing (NLP). Several syntactic analysis tasks adopt DGs as a formal model, given their efficiency in binary structure representation and syntactic disambiguation (Nivre, 2005; Kübler et al., 2009). Projects such as Universal Dependencies (De Marneffe et al., 2021) have consolidated DGs as the dominant paradigm in multilingual annotation systems, and recent works such as (Lopes et al., 2024; Souza et al., 2024; Duran & Pardo, 2024) demonstrate their ongoing relevance in current research. In this context, the present article aims to introduce the core formal foundations of DGs — connectivity, directionality, hierarchy, and syntac-

tic labelling — by discussing structural definitions of dependency, theoretical principles shared across different models, and finally, demonstrating their application through a simplified version of the Hays and Gaifman Grammar (Hays, 1964; Gaifman, 1965). The scope of this work is limited to the theoretical and abstract characterization of the essential properties of DGs, intending to provide an introductory yet systematic foundation for researchers interested in this approach.

### Keywords

syntax; dependency grammar

## 1. Introdução

As Gramáticas de Dependência<sup>1</sup>, doravante GD, têm ganhado cada vez mais relevância nos estudos linguísticos e computacionais, especialmente com o crescimento de pesquisas voltadas ao uso de sua metodologia para análise sintática automática. Esse modelo apresenta vantagens significativas, destacando-se pela sua capacidade de lidar de forma eficaz com a desambiguação sintática, graças à disposição binária das árvores que organiza os elementos da sentença em relações hierárquicas claras. Além disso, elas são muito mais eficientes para a realização de *parsing* automático, como apontado por Nivre (2005).

Como destaque, as *Universal Dependencies* (De Marneffe et al., 2021) são, no momento da escrita deste artigo, o maior expoente para as GD. Surgida em 2014, trata-se de uma proposta multilíngue de anotação sintática, ou seja, propõe diretrizes de anotação que sejam aplicáveis a diversas línguas. Isso possibilita a construção de modelos de processamento de linguagem natural que são mais robustos e generalizáveis, além de facilitar a comparação e análise entre as estruturas gramaticais das diversas línguas. Entretanto, mesmo que as *Universal Dependencies* se-

<sup>1</sup>O presente artigo é uma versão ampliada do texto disponível em *preprint* com DOI <https://dx.doi.org/10.22541/au.166906434.46213568/v1>

jam o tópico de maior destaque para as GD atualmente, este trabalho terá um foco mais teórico, voltado para o funcionamento abstrato dessas gramáticas. Assim, entende-se aqui que, dado o interesse atual por essa metodologia de análise sintática, um sobrevoo sobre os conceitos fundamentais pode ser interessante para quem quer se aprofundar no tema.

Considerando esse cenário, este artigo busca oferecer uma sistematização teórica acessível em língua portuguesa dos princípios formais que fundamentam as GD, como conectividade, direcionalidade, hierarquia e rotulação das relações sintáticas. Tal abordagem visa preencher uma lacuna na literatura em português, ao mesmo tempo em que oferece uma base conceitual sólida para pesquisadores interessados em aplicar ou desenvolver modelos baseados em dependência em contextos linguísticos e computacionais.

Além de seu interesse teórico, o modelo de dependência tem sido amplamente adotado em tarefas atuais de Processamento de Linguagem Natural, reforçando sua relevância científica. Ferramentas como as descritas por Nivre (2006) e De Marneffe et al. (2021) demonstram a eficácia das GD em sistemas de análise sintática automática, com ênfase na robustez e na generalização dos modelos treinados sobre corpora anotados em *Universal Dependencies*. Essa aplicabilidade é comprovada por diversos estudos recentes focados no português, como Miranda Jr. (2022), Duran & Pardo (2024), Lopes et al. (2024), Souza et al. (2024), Franco (2024), Pagani (2024) e Pagano et al. (2024), que mostram como as GD têm sustentado avanços em tarefas como desambiguação morfossintática, anotação automática e análise baseada em aspectos. Dessa forma, o presente trabalho se insere em uma tradição consolidada de uso computacional das GD, ao mesmo tempo em que busca contribuir com uma sistematização teórica acessível para esse modelo gramatical.

De fato, o conceito de dependência em língua natural é extremamente produtivo entre as diferentes formas de análise. Se tomados individualmente, cada estrato linguístico pode ser analisado em função de dependências. Fonética, morfologia, sintaxe, semântica — cada um desses níveis exprime alguma relação de dependência dentro de sua estrutura (Enfield, 2017). Todavia, as formas mais recorrentes de análises fundamentadas em dependência são sintáticas, recebendo o título de Gramática de Dependência.

Essas formas de estruturação textual podem ser recuperadas em textos antigos, como as gramáticas do Sânscrito de Panini, como também em trabalhos ao longo do século XX, como Sgall et al. (1986), Hellwig (1986), Mel'čuk (1988) e entre muitos outros. Mais recentemente, com o exponencial aumento de parsers utilizando GD como metodologia para o processamento de linguagem, elas ganharam um enorme destaque dentro do Processamento de Língua Natural (PLN) (Kübler et al., 2009; De Marneffe & Nivre, 2019).

Como destaque para os trabalhos de análise sintática baseados em dependências, pode-se citar o célebre trabalho de Lucien Tesnière “*Eléments de Syntaxe Structurale*”. Publicado postumamente em 1959, o livro de Tesnière é considerado o marco inicial dos estudos contemporâneos das GD. Nele, Tesnière define conceitos como *árvores*, *núcleos* e *nós* que são de extrema importância para as análises sintáticas com base em dependências posteriores aos seus trabalhos.

De maneira bem sucinta, pode-se definir o conceito de dependência em função das palavras pertencentes a um enunciado: “em uma sentença, todas, exceto uma palavra, dependem de outra(s) palavra(s)” (Debusmann, 2000). Ou seja, existem relações entre as palavras que constituem a sentença, de maneira que cada palavra necessita de outra, com exceção de apenas uma, para que esteja presente na sentença. No entanto, essa definição não diz muito sobre como identificar ou como construir essas relações entre os elementos de um enunciado.

A fim de elucidar o conceito de dependência e sua interação com as GD, este trabalho tem como objetivo oferecer uma base teórica acessível para pesquisadores que desejam compreender as GD como uma metodologia de análise sintática completa e autônoma, e não apenas como uma parte estrita no processamento computacional. Para isso, este artigo está organizado da seguinte forma: na primeira seção, será apresentada uma definição formal da relação de dependência entre os elementos de um enunciado; na segunda seção, será discutido como o conceito de dependência é empregado pelas GD na construção de sistemas gramaticais; por fim, na terceira seção, será demonstrado, por meio de uma versão simplificada da Gramática de Hays e Gaifman (Hays, 1964; Gaifman, 1965), como utilizar o conceito de dependência para a análise sintática de enunciados.

## 2. Dependência

Como trazido na introdução, pela definição de [Debusmann \(2000\)](#), as relações entre os elementos de um enunciado<sup>2</sup> recebem o nome de *dependência*, ou seja, entre dois elementos **a** e **b** relacionados em um enunciado, ou **a** depende de **b** ou **b** depende de **a**. Porém, saber que essas relações são dadas por dependência não é suficiente para definir, ou mesmo para identificar, as relações entre os elementos dentro de um enunciado.

Em [De Marneffe & Nivre \(2019\)](#) é possível encontrar uma lista de critérios para a identificação de dependência com base em [Zwicky \(2006 \[1985\]\)](#) e [Hudson \(1990\)](#), na qual a identificação é feita seguindo uma série de critérios em função da necessidade de processamento computacional, como também em função de diversas características linguísticas.

Segundo [De Marneffe & Nivre \(2019\)](#), a relação de dependência entre a cabeça (**H**), o dependente (**D**) e a construção (**C**) pode ser estabelecida com base nos seguintes critérios, conforme as propostas de [Zwicky \(2006 \[1985\]\)](#) e [Hudson \(1990\)](#):

- *H* determina a categoria sintática de *C* e normalmente pode substituir *C*;
- *H* determina a categoria semântica de *C*; *D* fornece a especificação semântica;
- *H* é obrigatória, *D* pode ser opcional;
- *H* seleciona *D* e determina quanto *D* é obrigatório ou opcional;
- A forma de *D* depende de *H*, em concordância ou governança;
- A posição linear de *D* é especificada com referência a *H*.

Mesmo com essa lista, ainda é complicado definir o conceito de *dependência*, uma vez que ela se utiliza de termos próprios das GD (*dependente*, *cabeça* e *governança*), por meio de características sintáticas e semânticas, para definir a relação de dependência, o que acaba tornando-a endógena, ou seja, ela explica *dependência* em função dos itens que fazem parte do conceito. Os próprios autores afirmam que é claro que esta lista contém uma mistura de critérios diferentes, alguns sintáticos e outros semânticos, e pode-se perguntar se existe uma única noção coerente

de dependência que corresponda a todos os diferentes critérios”<sup>3</sup> [De Marneffe & Nivre \(2019\)](#), p. 203).

Pretende-se aqui, por meio de uma definição não endógena, elucidar o conceito de dependência a fim de orientar o leitor para uma compreensão “unificada” deste conceito. Para isso, será utilizada a definição a presente no trabalho de [Robinson \(1970, p. 260\)](#), onde ela postula quatro axiomas para estruturas de dependência bem formadas com base nos trabalhos de [Hays \(1964\)](#) e [Gaifman \(1965\)](#) que formalizaram a proposta de [Tesnière \(2015 \[1959\]\)](#). Os axiomas são os seguintes:

- a* Um, e apenas um, elemento é independente;
- b* Todos os outros elementos dependem diretamente de algum outro elemento;
- c* Nenhum elemento depende diretamente de mais de um outro elemento;
- d* Se um elemento  $p_1$  depende diretamente do elemento  $p_2$  e existe um elemento  $p_3$  entre eles (na ordem linear da sequência de elementos), então  $p_3$  depende diretamente de  $p_1$ , de  $p_2$  ou de algum outro elemento entre eles.

Mais claramente: o axioma *a* define que só existe um elemento na sentença que não depende de nenhum outro; o axioma *b* estabelece que todos os demais elementos devem depender de algum outro; o axioma *c* define que nenhum elemento irá depender de mais de um elemento presente na sentença; e, por fim, o axioma *d* garante que, se uma palavra depende de outra e há uma terceira entre elas na frase, essa terceira não pode ficar isolada — ela deve depender de uma das duas ou de outra palavra intermediária, assegurando a coesão estrutural da análise de dependência.

Segundo [Robinson \(1970, p. 260\)](#), a relação *a depende de b* ( $\langle a, b \rangle \in R$ ) é transitiva, ou seja, *se a depende de b* e *b depende de c*, então *a indiretamente depende de c*; assimétrica, *a depende de b* e *b não depende de a*; e irreflexiva, não existe nenhum elemento *a* de tal que *a depende de a*, mais claramente, nenhum elemento pode depender dele mesmo.

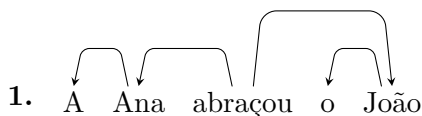
Pode-se agora, partindo dos axiomas, definir os termos *dependente*, *cabeça* e *governança* que fazem parte do conceito de dependência e estão presentes na definição de [De Marneffe & Nivre \(2019\)](#).

<sup>2</sup>Para evitar a ambiguidade envolvida com o termo palavra, utilizaremos, assim como [Debusmann \(2000\)](#), o conceito de elemento do enunciado, ou seja, cada um dos itens linguísticos que constituem o enunciado

<sup>3</sup>No original: *It is clear that this list contains a mix of different criteria, some syntactic and some semantic, and one may ask whether there is a single coherent notion of dependency corresponding to all the different criteria.*

Assim, como a relação *a depende de b* é binária, é necessário que existam dois elementos que a constituam. O elemento que *depende* recebe o nome de *dependente*, enquanto o elemento que *domina* a relação recebe o nome de *governador* ou *cabeça*.

Com essa nomenclatura pode-se definir a relação oposta a *dependência*, a *governança*, *a governa b* ( $\langle p_2, p_1 \rangle \in G$ ), de modo que ela é *transitiva*, se *a governa b* e *b governa c*, então *a indiretamente governa c*; *assimétrica*, se *a governa b*, então *b não governa a*; e *irreflexiva*, não existe nenhum elemento *a*, tal que *a governa a*, mais claramente, nenhum elemento pode governar ele mesmo. Porém, a *governança* permite que um elemento governe mais de um dependente, ou seja, um elemento que governa pode ter diversos dependentes, o que não ocorre na relação de *dependência*.



O exemplo 1. traz as relações de dependência na sentença: *A Ana abraçou o João*. No exemplo, as setas indicam o dependente, logo, o elemento ao lado oposto do arco representa o governante. Assim, podemos inferir que o conjunto  $\mathbf{R}$  que descreve a relação de dependência é constituído pelas tuplas:

2.  $\mathbf{R} = \langle A, Ana \rangle, \langle Ana, abraçou \rangle, \langle João, abraçou \rangle$  e  $\langle o, João \rangle$

Enquanto o conjunto  $G$ , que descreve a relação de governança, é constituído pelas tuplas:

3.  $\mathbf{G} = \langle Ana, A \rangle, \langle abraçou, Ana \rangle, \langle abraçou, João \rangle$  e  $\langle João, o \rangle$

Uma questão que merece atenção neste ponto é a relação entre itens funcionais — elementos cujo significado depende do funcionamento interno do sistema linguístico — e itens de conteúdo ou lexicais, que representam entidades externas a esse sistema. No exemplo 1., os artigos *a* e *o* são apresentados como dependentes dos nomes *Ana* e *João*, respectivamente. Essa escolha direcional, entre qual elemento governa e qual é governado, poderia ser feita também de forma inversa, sem prejuízo para a análise local da estrutura. A decisão de considerar os itens de conteúdo como governantes reflete apenas uma convenção adotada neste trabalho. Discussões mais amplas sobre a precedência estrutural entre

itens funcionais e itens de conteúdo envolvem escolhas teóricas específicas e extrapolam o escopo deste trabalho.

Em 1. há apenas as relações  $\langle a, b \rangle \in \mathbf{R}$ , ou seja, as dependências entre os pares de elementos do enunciado. As relações de dependência entre *a* e *b* não têm nenhuma distinção entre elas, o que não reflete completamente o sistema linguístico. Mesmo com a definição do conceito de *dependência*, não é possível realizar a descrição total do funcionamento de uma língua, tornando o conceito necessário, mas não suficiente para a descrição linguística por meio de GD.

A próxima seção pretende utilizar a definição de dependência aqui elencada para exemplificar como a maior parte das teorias utiliza esse conceito para a construção de descrições linguísticas com gramáticas de dependência.

### 3. Gramáticas de Dependência

Existem muitos trabalhos que se utilizam do conceito de dependência para realizar descrições linguísticas, essa diversidade gerou várias teorias gramaticais, como por exemplo a Gramática de Hays e Gaifman (Hays (1964); Gaifman (1965), a Descrição Gerativa Funcional (Functional Generative Description) (Sgall et al., 1986), a Gramática de Unificação de Dependências (Dependency Unification Grammar) (Hellwig, 1986), a Gramática de Palavras (Word Grammar) (Hudson, 1990), a Teoria do Texto-Sentido (Meaning-Text Theory) (Mel'čuk, 1988), e a Gramática de Dependência Funcional (Functional Dependency Grammar) (Järvinen & Tapanainen, 1998). Entretanto, ainda que muito numerosas e diversificadas, todas as teorias compartilham conceitos básicos para além do conceito de dependência.

Antes de apresentar uma teoria específica, como a Gramática de Hays e Gaifman (Hays, 1964; Gaifman, 1965), propomos aqui uma caracterização abstrata das propriedades que definem o núcleo conceitual das gramáticas de dependência. Essa abordagem metateórica permite compreender o que diferentes modelos compartilham, mesmo quando suas implementações analíticas divergem em relação a fenômenos como coordenação, translação ou não projetividade.

É possível encontrar em Mel'čuk & Polguère (2009, p. XIV) uma lista dos princípios básicos compartilhados pela maior parte, se não por todas, das teorias gramaticais fundamentadas no conceito de dependência, os quais são:

- i Conectividade da estrutura sintática;
- ii Direção das relações sintáticas;

- iii Organização hierárquica estrita da estrutura sintática;
- iv “Significado” das relações sintáticas.

A *conectividade da estrutura sintática* refere-se às relações entre os elementos de um enunciado. “A estrutura sintática forma um todo organizado, isto é, um sistema contínuo de relações sintáticas” (Mel’čuk & Polguère, 2009, p. XIV), ou seja, todo e qualquer elemento **a** existente na sentença se relaciona com ao menos um outro elemento **b**. Todos os elementos presentes em uma sentença fazem parte da estrutura sintática, de forma que nenhum fique de fora. Cada tupla  $\langle a, b \rangle \in \mathbf{R}$  representa uma frase mínima<sup>4</sup>, ou seja, um conjunto de dois elementos dotados de sentido.

Por exemplo, em 2., todos os elementos presentes na sentença de 1. fazem parte de ao menos uma das tuplas pertencentes a **R**. Também é verificável em 2. que nenhum elemento presente na sentença, em 1., está fora do conjunto **R**.

Mel’čuk & Polguère (2009) definem que a propriedade que se segue desse princípio é que “a estrutura sintática é um grafo conectado”, ou seja, a estrutura sintática, em função de dependências, é um conjunto relacional entre eles, permitindo que todos os elementos constituintes de um enunciado possam ser analisados como parte de um conjunto finito.

Antes de apresentar os princípios formais das GD, é útil estabelecer uma breve ponte com a teoria de grafos, dado que essa área da matemática fornece o aparato conceitual necessário para compreender a estrutura subjacente à análise sintática baseada em dependências.

De forma geral, um **grafo** é uma estrutura composta por um conjunto de nós (ou vértices) e conexões entre eles chamadas de arestas (ou arcos) (Diestel, 2017). Quando essas conexões possuem orientação — isto é, vão de um nó a outro com direção definida — temos um **grafo direcionado**. No contexto das GD, cada elemento de um enunciado (como uma palavra) pode ser representado por um nó, e a relação de dependência entre dois elementos por uma aresta direcionada, geralmente do dependente para o seu governante.

Continuando com as propriedades dos princípios, assim como a relação de dependência é assimétrica, as relações sintáticas estabelecidas por ela também serão, fazendo com que

as relações sintáticas sejam direcionadas, ou seja, partem de um elemento na direção de outro. É possível demonstrar o princípio *ii* pela capacidade de uma frase mínima expressar seu significado em função da categoria gramatical de seu governante. Tomemos os exemplos:

- 4. (a) Ana tem olhos **muito belos**  
(b) Ana tem olhos **castanhos**
- 5. (a) Ana corre **rápido**  
(b) Ana corre **muito bem**

Mesmo que exista um advérbio dentro da frase mínima *muito belo(s)*, em 4.a, ele não exprime comportamento adverbial, a frase mínima se comporta como um adjetivo, assim como *castanhos* em 4.b. Em 5.b, a frase mínima *muito bem* se comporta como um advérbio, assim como *rápido* em 5.a, uma vez que o governante do par é *bem*, um advérbio.

Mel’čuk & Polguère (2009, p. XIV) definem que a propriedade formal que segue do princípio de direcionalidade das relações sintáticas é que: a estrutura sintática é um grafo *conectado* e *direcionado*. Esse princípio faz com que a estrutura sintática seja um grafo conectado e organizado, existe um sentido e uma direção dentro do grafo, permitindo identificar quais são os elementos mais relevantes para a estrutura.

O princípio *iii* define que, assim como proposto pelos axiomas *a* e *b* de Robinson (1970, p. 260), cada elemento dentro da estrutura sintática é governado por um outro elemento, com exceção de um único elemento que governa todo enunciado. A organização hierárquica estrita define que cada frase mínima tenha um, e apenas um, governador e que exista um elemento máximo da estrutura de onde todos os outros elementos dependem, de maneira que o elemento que não é governado por nenhum outro é o item mais alto da estrutura de dependência.

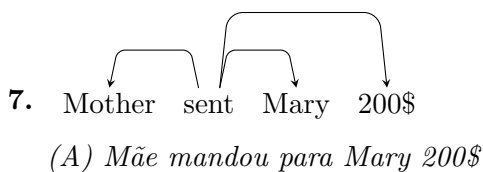
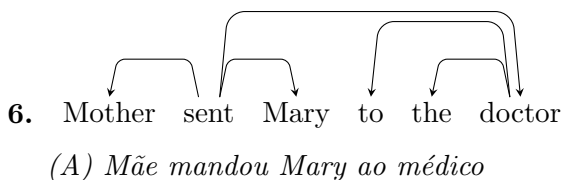
O princípio *iii* faz com que não seja possível a existência de ciclos infinitos no grafo da estrutura sintática. A estrutura parte do governador mais alto, passando pelos pares intermediários e termina nos dependentes que não governam nenhum elemento. Dessa forma, nenhum outro elemento irá dominar o governador máximo. Assim, no exemplo 1., o governador máximo do enunciado é o verbo *abraçar*, que não depende de nenhum outro elemento dentro da estrutura; os nomes *Ana* e *João* governam, respectivamente, os artigos *a* e *o*, que, por sua vez, não têm dependentes.

A propriedade formal que segue do princípio da organização hierárquica restrita do enunciado é que: “a estrutura sintática é um grafo

<sup>4</sup>Do inglês “*minimal phrase*”, a menor unidade sintática dotada de sentido (*o pedro, do joão, com você*). Adotamos tal notação pois os termos *sintagma* e *enunciado* como tradução de *phrase* já são marcados como termos de outras teorias

direcionado, acíclico e conectado, ou seja, uma árvore hierarquizada ou, mais brevemente, uma árvore” (Mel’čuk & Polguère, 2009, p. XV). Assim, segue-se do princípio *iii* que a estrutura sintática é um sistema télico, ou seja, ela parte de um elemento, o governador máximo, e termina em um ou mais outro(s) elemento(s).

Pelos princípios *i*, *ii* e *iii*, a estrutura sintática ainda é um reflexo das relações de dependência, porém, como já dito, apenas as dependências não são suficientes para exprimir a sintaxe das línguas naturais. Um exemplo disso é a possibilidade de haver duas relações de dependência *R1* e *R2* constituídas pelos mesmos elementos **a** e **b**, mas que exprimem relações sintáticas distintas. Para exemplificar, tomemos as sentenças *Mother Sent Mary to the doctor* e *Mother sent Mary 200\$* trazidas por Mel’čuk & Polguère (2009, p. XV):



Em ambos os exemplos, observa-se o mesmo par relacional  $\langle \text{Mary, to send} \rangle \in \mathbf{R}$ . Ainda que, em ambas as ocorrências, *Mary* estabeleça uma relação de dependência com *to send*, tal configuração exprime, em cada caso, uma relação sintática distinta.

Em 6., *Mary* e *to send* se relacionam por meio de um caso acusativo<sup>5</sup>, enquanto em 7., os mesmos elementos estão vinculados por um caso dativo<sup>6</sup>. Ainda que o par relacional  $\langle \text{Mary, to send} \rangle \in \mathbf{R}$  possua a mesma distribuição linear nos enunciados — com *Mary* ocorrendo logo após *to send* —, a interpretação sintática atribuída à relação varia entre os contextos, refletindo funções gramaticais distintas.

Dessa forma, as relações de dependência, para que sejam coesas, não se limitam a uma mera distribuição estrutural; elas também carregam um valor semântico. Cabe destacar, contudo, que

<sup>5</sup>O caso acusativo marca, em geral, o objeto direto da ação verbal, ou seja, o participante afetado diretamente pelo evento (Blake, 2001, cap. 5).

<sup>6</sup>O caso dativo é tradicionalmente associado ao objeto indireto, representando o destinatário, beneficiário ou meta da ação verbal (Blake, 2001, cap. 5).

esse valor não é fixo ou previamente determinado; ele pode variar conforme a natureza específica do elemento governante da frase mínima, como se pode observar nos exemplos a seguir.

8. A UFSCar contratou novos professores

9. A UFSCar sofreu cortes do governo

10. A UFSCar sediará um evento internacional este ano

Mesmo que a relação  $\langle \mathbf{a}, \mathbf{b} \rangle \in \mathbf{R}$  entre *UFSCar* e os verbos das sentenças indique sintaticamente sujeito, em cada exemplo *UFSCar* desempenha papéis semânticos distintos: AGENTE, PACIENTE e LOCATIVO<sup>7</sup>.

Finalizando as propriedades que seguem dos princípios, a propriedade formal que segue do princípio *iv* é que: “a estrutura sintática é uma árvore em que os nós são rotulados e vinculados por arcos nomeados pelas relações sintáticas que representam” (Mel’čuk & Polguère, 2009). Assim, cada arco que liga os elementos do grafo da estrutura terá uma nomenclatura específica para representar sua função sintática.

Os princípios *i*, *ii*, *iii* e *iv* abarcam a maior parte das teorias gramaticais de dependência, mas eles ainda não são suficientes para solucionar problemas cotidianos da linguagem. Um dos problemas mais evidentes é a coordenação. Tomemos os exemplos:

11. Pedro e João visitaram a praia

12. Pedro gosta de cães e gatos

13. Pedro comprou e comeu um bolo

Pelo princípio *iii*, as relações sintáticas devem ser estritamente hierarquizadas, todavia isso é contraintuitivo nos exemplos. Tanto *Pedro* quanto *João* são ótimos candidatos para dependente na relação de sujeito do verbo *visitar*, em 11., assim como *cães* e *gatos* são ótimos candidatos para complemento de *gostar*, em 12., e *comprar* e *comer* são ótimos candidatos para governantes em 13.. O conceito de dependência, em princípio, não é suficiente para esse tipo de relação.

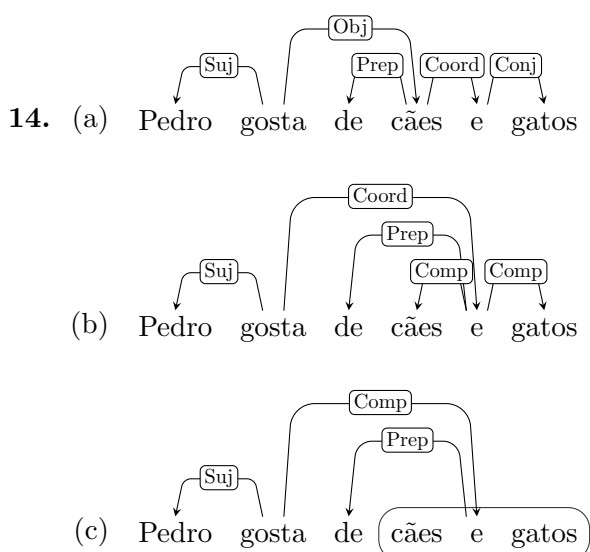
<sup>7</sup>Papéis semânticos são as funções que os participantes de um evento desempenham em relação ao predicado, como Agente, Paciente e Locativo. Essas funções derivam de propriedades semânticas associadas aos argumentos: AGENTE – entidade animada que inicia ou controla a ação expressa pelo predicado, típica do sujeito de verbos transitivos; PACIENTE – participante afetado pela ação, que sofre alteração ou movimento causado pelo evento verbal; LOCATIVO – argumento que indica o local onde ocorre o evento, podendo ser exigido pelo verbo ou introduzido como adjunto (Cançado, 2005, p. 66-71).

Diversas soluções foram propostas para esse problema. Em [Ternière \(2015 \[1959\]\)](#), além da relação de *dependência*, são definidas outras duas relações sintáticas fundamentais. A primeira é a *junção* (do francês *jonction*), responsável pelas estruturas de coordenação, como ilustram os exemplos [11.](#), [12.](#) e [13.](#). A segunda é a *translação* (do francês *translation*), que tem a função de alterar a categoria lexical de um elemento. Um exemplo de *translação* pode ser observado na sequência *de frango*, em *estrogonofe de frango*: o uso da preposição *de* permite que o nome *frango* modifique o nome *estrogonofe*, papel que, nas construções mais prototípicas, seria atribuído à relação de dependência com adjetivos.

Outras soluções possíveis para o problema da coordenação podem ser observadas no exemplo [14.](#), abaixo. Para [14.a](#), tem-se a proposta de [Mel'čuk \(1988\)](#), em que a coordenação é tratada como uma relação de dependência, onde o primeiro elemento da coordenação é o governante de quem a conjunção depende, que, por sua vez, governa o último elemento da coordenação.

Em [14.b](#), tem-se a proposta de [Sgall et al. \(1986\)](#), na qual a relação de dependência *coord* permite que a conjunção *e* seja o governante da estrutura de coordenação, fazendo com que os elementos da coordenação sejam dependentes dela.

E no exemplo [14.c](#), tem-se a proposta de [Kahane \(1997\)](#), nela as relações de dependência não ocorrem apenas entre os elementos da sentença, mas podem ocorrer em um conjunto de elementos (que ele chama de *bolhas*<sup>8</sup>), fazendo com que toda a estrutura *cães e gatos* seja apenas uma *bolha*.

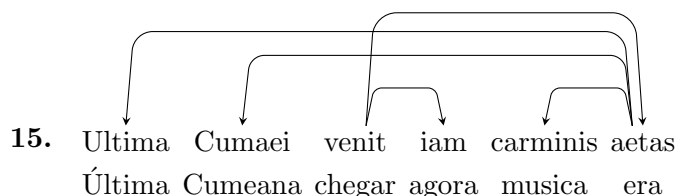


<sup>8</sup>Do inglês “bubbles”

Em todas as propostas, a hierarquia estrita entre os elementos é preservada; contudo, cada teoria adota uma estratégia distinta para operacionalizar esse princípio. Por se tratar de um texto introdutório, voltado à explicação dos fundamentos das GD, não abordaremos em profundidade as diferentes soluções propostas para o tratamento de estruturas coordenadas. Além disso, essas estruturas não serão contempladas na seção seguinte, cujo objetivo é apenas exemplificar o funcionamento reduzido de uma gramática de dependências.

Por fim, um outro conceito relevante para as teorias gramaticais com base em dependência é a *projetividade* ([Lecerf, 1960](#); [Hays, 1964](#); [Gaifman, 1965](#)), isto é, para que uma árvore de dependências seja projetiva, os galhos não se intersectam, ou seja, as dependências da árvore não se cruzam entre elas. Ainda que a maior parte das sentenças da maioria das línguas naturais seja projetiva ([Mel'čuk & Polguère, 2009](#), p. 85), é possível encontrar exemplos em que a *projetividade* é quebrada.

Como exemplo, línguas que marcam caso morfológicamente, como o latim, o russo ou o húngaro, aceitam distribuições livres entre os elementos sintáticos, uma vez que a função sintática é demarcada pela morfologia e não pela ordem. Essa liberdade com relação à ordenação faz com que existam sentenças que violem a *projetividade*.



O exemplo [15.](#) é um excerto da quarta écloga de Virgílio trazido por [Mel'čuk & Polguère \(2009](#), p. 87). Nele, temos a sentença “É chegada agora a era da última canção cumeana” em latim, na qual é possível observar que o arco que representa a dependência entre *aetas* e *venit* intersecta (cruza) os arcos que representam as dependências de *Ultima* e *Cumaei* em relação a *aetas*.

As teorias de descrição gramatical tratam a falta de projetividade de diferentes maneiras. Em sua maioria, elas propõem soluções multi-estrato. Nas propostas de [Sgall et al. \(1986\)](#), [Mel'čuk \(1988\)](#) e [Hudson \(1990\)](#) a estrutura de dependência não está associada apenas à sintaxe, mas também a outros níveis de granularidade linguística, como morfologia e semântica.

Dentro das teorias multiestrato, algumas, como Sgall et al. (1986), assumem que a falta de projetividade só existe nos níveis mais altos, como a sintaxe, e é solucionada nos níveis mais profundos, como a semântica. Outras propostas, como as de Mel'čuk (1988) e Hudson (1990), recorrem a transformações lógicas e interações entre os diferentes níveis para solucionar a falta de projetividade.

Assim como no caso das estruturas coordenadas, optamos aqui por adotar uma abordagem simplificada. Embora seja possível encontrar sentenças não projetivas em línguas naturais, a maior parte das sentenças efetivamente o é; logo, adotaremos a projetividade como padrão neste trabalho por motivos de simplificação metodológica. Essa decisão visa facilitar a exposição dos fundamentos das Gramáticas de Dependência, já que estruturas projetivas evitam cruzamentos entre dependências e tornam os exemplos mais transparentes. Ainda assim, é importante destacar que fenômenos de não projetividade ocorrem em todas as línguas naturais e que há uma ampla discussão na literatura sobre como modelá-los adequadamente<sup>9</sup>.

Finalizada aqui, esta seção trouxe o levantamento das características comuns que regem as teorias gramaticais baseadas em dependência. Toda a discussão se estabeleceu de maneira abstrata para que fosse possível relacionar os conceitos com diferentes teorias, mas em nenhum momento foi abordada uma teoria em específico. Para que seja possível observar o funcionamento dos princípios e propriedades levantados ao longo deste artigo, a próxima seção apresentará o funcionamento do modelo proposto por Hays (1964) e Gaifman (1965), um modelo simples, mas que é de grande influência em diversos trabalhos sobre dependência (Lai & Huang, 1998; Lombardo & Lesmo, 1998; Tapanainen & Järvinen, 1997; Nivre & Kahane, 2004; Yli-Jyrä, 2005; Gómez-Rodríguez et al., 2022, entre outros).

#### 4. Gramática de Dependência de Hays e Gaifman

Para demonstração do comportamento de um sistema de uma gramática de dependências, examinaremos a proposta da Gramática de Hays e Gaifman (Hays, 1964; Gaifman, 1965), doravante GHG, uma vez que, além de ser a base dos postulados de Robinson (1970) para estruturas de dependência bem formadas, ela influenciou teo-

rias como as de Lombardo & Lesmo (1998) e Lai & Huang (1998) e também pode ser recuperada em trabalhos recentes como Morey et al. (2018).

De fato, Hays (1964) e Gaifman (1965) são os primeiros autores a propor um sistema de axiomas necessário e suficiente para a implementação de uma gramática de dependência em algoritmos autônomos de análise textual. Ambos os trabalhos consistem em reestruturações formais da proposta de Tesnière (2015 [1959]), mas deslocam o foco da descrição linguística para uma modelagem estritamente formal da sintaxe, centrada exclusivamente na relação de *dependência*, deixando de lado as relações de *junção* e *translação* originalmente descritas por Tesnière.

Essa decisão metodológica decorre da orientação algébrica e computacional que norteia as propostas desses autores. Em contextos de formalização sintática voltados para aplicações automáticas, a relação de dependência — por sua natureza binária, hierárquica e orientada — oferece maior compatibilidade com esquemas lógicos e algoritmos eficientes. Por outro lado, as relações de junção e translação, mais associadas a fenômenos como coordenação e mudança categorial, exigiriam um aparato teórico adicional que dificultaria a operacionalização computacional pretendida nos modelos de Hays (1964) e Gaifman (1965).

Para os fins deste artigo, que visa apresentar de forma introdutória o funcionamento de uma gramática de dependências, tal simplificação mostra-se inclusive vantajosa. Ainda que não contemple diretamente as estruturas de junção e translação, o modelo adotado é suficiente para explicitar os princípios fundamentais que regem esse tipo de abordagem sintática. A exclusão dessas relações, portanto, não compromete a proposta didática aqui assumida, permitindo que se evidenciem, de maneira clara e gradual, os mecanismos centrais que estruturam uma gramática de dependência.

Em continuidade à proposta formal, Gaifman (1965, p. 305–306) define uma gramática de dependência como composta por três conjuntos principais de regras, responsáveis por estruturar como as palavras de uma sentença se relacionam:

- L1** Regras da forma  $X(Y_1, \dots, Y_m * Y_{m+1}, \dots, Y_n)$ , nas quais  $X$  ocupa a posição do asterisco (\*) e representa a categoria principal. As categorias  $Y_1$  até  $Y_n$  são os possíveis dependentes que podem aparecer ao redor de  $X$ , respeitando a ordem indicada. O valor de  $m$  pode ser igual a 0 e/ou igual a  $n$ , o que permite que  $X$  apareça sozinho ou com qualquer número de dependentes.

<sup>9</sup>Para além dos trabalhos já citados, o leitor também pode consultar os trabalhos de Tapanainen & Järvinen (1997) e Kahane et al. (1998).

- LII** Regras que listam, para cada categoria  $X$ , quais elementos (palavras) pertencem a ela. Um mesmo elemento pode pertencer a mais de uma categoria.
- LIII** Regra que define quais categorias podem servir como núcleo de uma sentença — isto é, aquelas que não precisam depender de nenhuma outra.

Em termos mais simples, **L<sub>I</sub>** trata da forma como as categorias se combinam dentro da sentença, ou seja, como as palavras se organizam e se ligam umas às outras. Veja os exemplos abaixo:

16. O João foi para o escritório

O João foi para o escritório  
**Det Nom Verb Prep Det Nom**

17. Pedro alcançou o ônibus

Pedro alcançou o ônibus  
**Nom Verb Det Nom**

Já **L<sub>II</sub>** trata da classificação das palavras em categorias gramaticais. Por exemplo:

18. Det = {O, A, Os, As, Um, Uma, Uns, Umas, etc.}

19. Nom = {Pedro, João, ônibus, bolo, escritório, Maria, etc.}

Por fim, **L<sub>III</sub>** define quais categorias podem aparecer como núcleo de uma sentença — ou seja, quais não precisam depender de outras. Normalmente, o verbo é o governante principal da sentença, mas há exceções, como enunciados nominais ou deverbais, que possibilitam outras categorias como núcleos.

Segundo Gaifman (1965, p. 306), uma sentença é composta por um conjunto de elementos —  $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_i$  —, sendo cada um associado a uma categoria —  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_i$  — com o mesmo índice. As palavras se ligam por relações binárias de dependência, representadas como  $\mathbf{d}(\mathbf{p}_i, \mathbf{p}_j)$ , indicando que  $\mathbf{p}_i$  depende de  $\mathbf{p}_j$ . A partir de  $\mathbf{d}$ , define-se  $\mathbf{d}^*$ , sua forma transitiva.

O sistema obedece às seguintes propriedades formais:

- I. Nenhum elemento pode depender de si mesmo, direta ou indiretamente:  $\mathbf{d}^*(\mathbf{p}_i, \mathbf{p}_i)$  não ocorre.
- II. Cada elemento da sentença depende de pelo menos um outro: existe  $\mathbf{d}(\mathbf{p}_i, \mathbf{p}_j)$  para cada  $\mathbf{p}_i$ .
- III. Se  $\mathbf{p}_i$  depende (direta ou indiretamente) de  $\mathbf{p}_j$  e houver um termo intermediário  $\mathbf{p}_k$  entre eles, então  $\mathbf{p}_k$  também depende (direta ou indiretamente) de  $\mathbf{p}_j$ .

IV. Todos os elementos da sentença devem estar interligados por relações de dependência.

V. Se elementos com categorias  $\mathbf{X}_1, \dots, \mathbf{X}_n$  dependem de um elemento com categoria  $\mathbf{X}$ , então existe uma regra **L<sub>I</sub>** do tipo  $\mathbf{X}(\mathbf{X}_1, \dots, *, \dots, \mathbf{X}_n)$ .

VI. O único elemento que não depende de nenhum outro — o governante da sentença — deve pertencer a uma categoria listada em **L<sub>III</sub>**.

Essas propriedades asseguram o funcionamento consistente do sistema. Vejamos de forma mais didática:

- **I** impede ciclos: nenhuma palavra pode depender de si mesma.
- **II** e **VI** garantem que todos os elementos, exceto o governante principal, dependam de outro.
- **III** assegura que a estrutura de dependência respeite uma hierarquia linear e coerente.
- **IV** garante que não haja elementos “isolados” na sentença.
- **V** conecta a estrutura sintática com as regras da gramática, validando a construção da sentença.
- **VI** restringe os possíveis governantes às categorias previamente autorizadas.

Desse modo, essas regras fornecem um modelo formal e sistemático para representar sentenças por meio de relações hierárquicas de dependência entre palavras.

As propriedades formais descritas acima articulam-se diretamente a três dos princípios estabelecidos por Mel'čuk & Polguère (2009), discutidos na Seção 3, bem como às condições propostas por Robinson (1970) para a definição da relação de dependência, discutidas na Seção 2. O princípio *i*, que exige a conectividade da estrutura sintática, é contemplado pela propriedade **IV**, segundo a qual todos os elementos da sentença devem estar conectados por meio de relações de dependência.

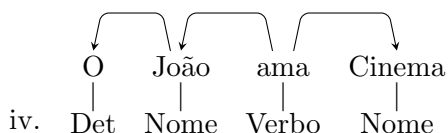
O princípio *ii*, por sua vez, diz respeito à direção obrigatória dessas relações, e encontra correspondência nas propriedades **II** e **VI**, que garantem que cada item dependa de um outro (com exceção do governante principal). Esse direcionamento reforça a assimetria e a orientação da estrutura, características também presentes na formulação de Robinson, que define a relação de dependência como binária, irreflexiva e direcionada.

Já o princípio *iii*, que postula uma organização hierárquica estrita entre os constituintes, está refletido nas propriedades **I** e **III**, responsáveis por impedir ciclos e garantir a transitividade das relações. Essas condições asseguram que a estrutura sintática funcione como um grafo hierarquicamente ordenado. O princípio *iv*, que trata do “significado” das relações sintáticas, será discutido mais à frente nesta seção, em conexão com os aspectos semânticos e informacionais da dependência.

Dessa forma, observa-se uma convergência entre os princípios sintáticos, a formalização lógica e as restrições linguísticas que fundamentam a abordagem de dependência como um modelo descritivo robusto e teoricamente consistente.

#### D1 Demonstração simplificada do funcionamento da GHG:

- i.  $O_1$  João<sub>2</sub> ama<sub>3</sub> cinema<sub>4</sub>
- ii. Verb<sub>3</sub>(Nome<sub>2</sub>, \* Nome<sub>4</sub>)  
Nome<sub>2</sub>(Det<sub>1</sub>, \*)  
Nome<sub>4</sub>(\*)  
Det<sub>1</sub>(\*)
- iii. Verb = {amar, ...}  
Nome = {João, cinema, ...}  
Det = {O, ...}



A demonstração **D1** evidencia, de forma simplificada, o funcionamento da GHG, permitindo observar, passo a passo, como os diferentes conjuntos que compõem o sistema interagem na construção da estrutura sintática. A sentença analisada em **D1-i** é segmentada em elementos com índices lineares, que serão utilizados para definir as dependências.

No conjunto **D1-ii**, as regras de ordenação categorial (**L<sub>I</sub>**) estabelecem como as categorias podem se combinar. O verbo *amar*, por exemplo, exige dois nomes como dependentes — um à esquerda (o sujeito *João*) e outro à direita (o objeto *cinema*). Por sua vez, o nome *João* admite um determinante (*O*) à esquerda. Esses esquemas de combinação refletem a estrutura sintática da sentença, organizando-a a partir de regras formais baseadas em categorias.

Já o conjunto **D1-iii** apresenta as regras de associação entre os itens lexicais e suas categorias, ou seja, o mapeamento dos elementos da sentença para os símbolos auxiliares que os classificam. Essa etapa vincula *amar* à categoria

**Verb**, *João* e *cinema* à categoria **Nome**, e *O* à categoria **Det**.

Por fim, a estrutura representada em **D1-iv** é o resultado da aplicação das regras anteriores. A árvore mostra que *amar* governa tanto *João* quanto *cinema*, enquanto *João* governa o determinante *O*. Trata-se de uma estrutura conectada, acíclica, hierarquizada e orientada, como esperado para um sistema baseado em dependências.

Assim, pode-se definir que, como proposto por Debusmann (2000, p. 7), uma gramática de dependência, como a GHG, é constituída de quatro conjuntos distintos, como na definição a seguir:

Gramática de Dependência de Hays e Gaifman  
( $\langle R, L, C, F \rangle$ )

Em que:

$R$  é o conjunto de regras de dependência entre as categorias sintáticas;  
 $L$  é o conjunto de símbolos terminais, ou elementos;  
 $C$  é o conjunto de símbolos auxiliares, ou categorias sintáticas; e  
 $F$  é uma função de atribuição de  $L$  para  $C$  ( $F: L \rightarrow C$ ), ou seja, uma regra de associação entre os elementos e suas categorias sintáticas.

A demonstração apresentada em **D1** permite visualizar, de maneira clara, o papel desempenhado por cada um desses componentes. As regras de combinação entre categorias **D1-ii** constituem o conjunto **R**; os itens lexicais da sentença **D1-i** formam o conjunto **L**; suas respectivas categorias **D1-iii** compõem o conjunto **C**; e a associação entre elementos e categorias é formalizada pela função **F**. A árvore de dependência final **D1-iv** é, portanto, a projeção dessas quatro camadas atuando em conjunto sobre uma sentença concreta.

O intuito desta seção é demonstrar o comportamento de uma gramática de dependências utilizando os princípios e propriedades apresentados anteriormente. Ainda que a GHG atenda diretamente aos princípios de conectividade, direção e hierarquia sintática, ela não contempla, em sua formulação original, o princípio *iv*, relativo ao “significado” das relações sintáticas.

Conforme discutido por Mel'čuk & Polguère (2009), esse princípio exige que cada relação de dependência seja dotada de um sentido funcional ou semântico, o que vai além da simples estruturação formal das dependências. A GHG, entretanto, apenas indica a existência de dependência entre categorias, sem especificar o papel desempenhado pelos elementos em suas relações.

Ainda assim, a nomenclatura das relações de dependência pode ser inferida a partir da distribuição categorial das dependências, o que permite a introdução de sentidos por meio de um conjunto adicional de regras.

Nesse sentido, propõe-se uma extensão formal da gramática, conforme Debusmann (2000, p. 7), com a adição de dois novos conjuntos:

Gramática de Dependência de Hays e Gaifman com a nomenclatura das relações sintáticas  $\langle R, N, L, C, F, G \rangle$

Em que:

$R, L, C$  e  $F$  são análogos à definição anterior;

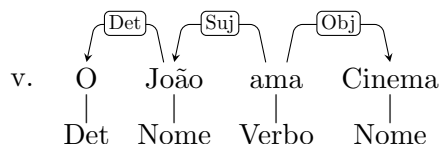
$N$  é o conjunto de sentidos atribuídos às regras de  $R$ ; e

$G$  é uma função de atribuição de  $R$  para  $N$  ( $G: R \rightarrow N$ ), ou seja, uma regra de associação entre as regras de dependência e seus sentidos sintáticos.

Essa reformulação permite que a demonstração **D1** seja estendida, resultando na demonstração **D2**, com a adição de um novo passo (iv), responsável por explicitar a nomenclatura das relações sintáticas:

## D2

- i.  $O_1$  João<sub>2</sub> ama<sub>3</sub> cinema<sub>4</sub>
- ii.  $\text{Verb}_3(\text{Nome}_2, * \text{Nome}_4)$   
 $\text{Nome}_2(\text{Det}_1, *)$   
 $\text{Nome}_4(*)$   
 $\text{Det}_1(*)$
- iii.  $\text{Verb} = \{\text{amar}, \dots\}$   
 $\text{Nome} = \{\text{João}, \text{cinema}, \dots\}$   
 $\text{Det} = \{\text{O}, \dots\}$
- iv. Sujeito =  $\text{Verb}(\text{Nome}, *)$   
 $\text{Det} = \text{Nome}(\text{Det}, *)$   
 $\text{Objeto} = \text{Verb}(*, \text{Nome})$



Os primeiros quatro passos da demonstração **D2** seguem diretamente os passos já descritos em **D1**, com a diferença de que, no passo **D2–iv**, introduz-se um mapeamento entre a estrutura categorial das dependências e os rótulos funcionais tradicionalmente atribuídos às relações sintáticas. Por exemplo, a regra  $\text{Verb}(\text{Nome}, *)$  é associada ao papel de *sujeito*,

enquanto  $\text{Verb}(*, \text{Nome})$  é associada a *objeto*. Esse vínculo entre estrutura categorial e sentido sintático é formalizado por meio da função  $G$ , como descrito na definição formal acima.

O passo **D2–v** apresenta a estrutura final de dependência, agora com os arcos devidamente rotulados segundo os sentidos sintáticos deduzidos. Esse modelo estendido satisfaz não apenas os três primeiros princípios descritos na seção anterior (conectividade, direção e hierarquia), mas também o princípio *iv*, ao incorporar o “significado” das relações. Dessa forma, a gramática  $\langle R, N, L, C, F, G \rangle$  constitui uma formulação mais completa e interpretável da GHG, compatível com os requisitos estruturais e semânticos das GD.

As demonstrações aqui apresentadas evidenciam como a GHG é capaz de representar, de forma sistemática, os princípios estruturais que regem as GD. A inclusão de uma camada interpretativa, como em **D2**, amplia esse modelo ao contemplar também o princípio do significado das relações sintáticas.

Ainda que simplificadas, essas demonstrações são suficientes para ilustrar, de modo introdutório, o funcionamento interno de uma gramática de dependências. Elas tornam visíveis os mecanismos formais que organizam a estrutura sintática dos enunciados, oferecendo uma base sólida para a compreensão de elementos fundamentais das GD.

## 5 Considerações Finais

Este artigo teve como objetivo apresentar, de maneira introdutória e formal, os fundamentos centrais das Gramáticas de Dependência (GD), enfatizando tanto o conceito de dependência quanto os princípios estruturais que regem sua aplicação em teorias linguísticas. Foram abordadas três dimensões principais: (i) a definição formal da relação de dependência, a partir do trabalho de Robinson (1970); (ii) a sistematização dos princípios estruturais comuns às principais GD, como proposto por Mel'čuk & Polguère (2009); e (iii) a demonstração, com base na Gramática de Hays e Gaifman (GHG) (Hays, 1964; Gaifman, 1965), de como esses princípios podem ser aplicados na construção de análises linguísticas.

As características essenciais abordadas neste trabalho incluem: a conectividade, a direcionalidade, a hierarquia e a semântica das relações sintáticas — propriedades que, quando formalizadas, permitem representar estruturas linguísticas como árvores sintáticas orientadas, acíclicas e rotuladas.

A escolha por não incluir, neste artigo, as relações de junção e translação — originalmente propostas por Tesnière (2015 [1959]) — decorre de uma decisão metodológica orientada pela própria estrutura da GHG, que prioriza a dependência como única relação sintática formalizada. Essa delimitação não compromete a generalidade da abordagem, mas permite evidenciar com maior precisão os princípios fundamentais das GD em sua formulação clássica, tal como empregada em diversos modelos computacionais. Ao concentrar-se na estrutura projetiva e nas dependências binárias hierárquicas, o artigo busca esclarecer os mecanismos essenciais de organização sintática, sem a sobrecarga de fenômenos que demandariam arcabouços teóricos adicionais.

A relevância do tema é reforçada pela ampla utilização das GD em tarefas contemporâneas de Processamento de Linguagem Natural (PLN), especialmente em contextos multilíngues como o projeto Universal Dependencies (De Marneffe et al., 2021). Estudos recentes apresentados em eventos como o STIL (Lopes et al., 2024; Souza et al., 2024) e trabalhos aplicados ao português brasileiro (Duran & Pardo, 2024; Franco, 2024; Pagani, 2024; Pagano et al., 2024) demonstram a atualidade e a vitalidade do campo, evidenciando a importância de consolidar, em português, revisões teóricas que possam servir de base à formação de novos pesquisadores.

Portanto, espera-se que este artigo contribua para tornar mais acessível o universo das GD, não apenas como técnica de anotação ou modelo computacional, mas como uma abordagem teórica completa para a descrição sintática das línguas naturais.

## Agradecimentos

Este trabalho foi realizado no âmbito do Centro de Inteligência Artificial da Universidade de São Paulo (C4AI -<http://c4ai.inova.usp.br/>), com o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (processo FAPESP 2019/07665-4) e da IBM. Este projeto também foi apoiado pelo Ministério da Ciência, Tecnologia e Inovações, com recursos da Lei N. 8.248, de 23 de outubro de 1991, no âmbito do PPI-Softex, coordenado pela Softex e publicado como Residência em TIC 13, DOU 01245.010222/2022-44.

Isaac Souza de Miranda Jr. recebeu apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, processo nº 23/01892-4).

Agradecemos também aos pareceristas Elvis de Sousa e Daniela Barreiro Claro pelas leitu-

ras atentas e pelas sugestões criteriosas. Seus comentários foram fundamentais para aprimorar a estrutura, a clareza e a fundamentação teórica do artigo.

## Referências

- Blake, Barry J. 2001. *Case*. Cambridge University Press. doi 10.1017/CB09781139164894
- Cançado, Márcia. 2005. Posições argumentais e propriedades semânticas. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada* 21(1). 23–56. doi 10.1590/S0102-44502005000100002
- De Marneffe, Marie-Catherine, Christopher D Manning, Joakim Nivre & Daniel Zeman. 2021. Universal dependencies. *Computational linguistics* 47(2). 255–308. doi 10.1162/coli\_a\_00402
- De Marneffe, Marie-Catherine & Joakim Nivre. 2019. Dependency grammar. *Annual Review of Linguistics* 5(1). 197–218. doi 10.1146/annurev-linguistics-011718-011842
- Debusmann, Ralph. 2000. An introduction to dependency grammar. *Hausarbeit für das Hauptseminar Dependenzgrammatik SoSe* 99(1). 1–16
- Diestel, Reinhard. 2017. *Graph theory*. Springer 5th edn. doi 10.1007/978-3-662-70107-2
- Duran, Magali Sanches & Thiago Alexandre Salgueiro Pardo. 2024. Anotação de corpus, um lugar privilegiado de observação linguística: um estudo das aposições do português brasileiro segundo o modelo universal dependencies. Em *Encontro de Linguística de Corpus (ELC)*, 118–123. ↗
- Enfield, Nick J. (ed.). 2017. *Dependencies in language: On the causal ontology of linguistic systems*. Language Science Press. doi 10.5281/zenodo.573773
- Franco, Francisco Dias. 2024. *ABSAuDA-Análise de sentimentos baseado em aspectos utilizando análise de dependência*. Universidade Federal de Pelotas. Tese de Mestrado
- Gaifman, Haim. 1965. Dependency systems and phrase-structure systems. *Information and Control* 8(3). 304–337. doi 10.1016/S0019-9958(65)90232-9
- Gómez-Rodríguez, Carlos, Morten H. Christensen & Ramon Ferrer-i Cancho. 2022. Memory limitations are hidden in grammar. *Glottometrics* 52. 39–64. doi 10.53482/2022\_52\_397

- Hays, David G. 1964. Dependency theory: A formalism and some observations. *Language* 40(4). 511–525. doi 10.2307/411934
- Hellwig, Peter. 1986. Dependency unification grammar. Em *11<sup>th</sup> International Conference on Computational Linguistics (COLING)*, 195–198. ↗
- Hudson, Richard A. 1990. *English word grammar*. Blackwell
- Järvinen, Timo & Pasi Tapanainen. 1998. Towards an implementable dependency grammar. Em *Dependency-Based Grammars*, ↗
- Kahane, Sylvain. 1997. Bubble trees and syntactic representations. Em *5<sup>th</sup> Meeting of Mathematics of Language (MOL5)*, 70–76. ↗
- Kahane, Sylvain, Alexis Nasr & Owen Rambow. 1998. Pseudo-projectivity, a polynomially parsable non-projective dependency grammar. Em *36<sup>th</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 646–652. doi 10.3115/980845.980953
- Kübler, Sandra, Ryan McDonald & Joakim Nivre. 2009. Dependency parsing. Em *Dependency parsing*, 11–20. Springer. doi 10.2200/S00169ED1V01Y200901HLT002
- Lai, Tom B. Y. & Changning Huang. 1998. Complements and adjuncts in dependency grammar parsing emulated by a constrained context-free grammar. Em *Processing of Dependency-based Grammars*, 102–108. ↗
- Lecerf, Yves. 1960. Programme des conflits, modèle des conflits. *Bulletin bimestriel de l'ATALA* 1(4). 11–18
- Lombardo, Vincenzo & Leonardo Lesmo. 1998. Unit coordination and gapping in dependency theory. Em *Dependency-based Grammars*, 11–20. ↗
- Lopes, Lucelene, Magali Sanches Duran & Thiago Alexandre Salgueiro Pardo. 2024. Desambiguação de lema e atributos morfológicos na anotação do corpus porttinari-base. Em *Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL)*, 336–345. doi 10.5753/stil.2024.245213
- Mel'čuk, Igor. 1988. *Dependency syntax: theory and practice*. SUNY press
- Mel'čuk, Igor & Alain Polguère. 2009. *Dependency in linguistic description*. De Gruyter Brill. doi 10.1075/slcs.111
- Miranda Jr., Isaac Souza de. 2022. *Não é nada não: uma análise das lexias negativas do português para anotação nas universal dependencies*. Universidade Federal de São Carlos. Tese de Mestrado. ↗
- Morey, Mathieu, Philippe Muller & Nicholas Asher. 2018. A dependency perspective on RST discourse parsing and evaluation. *Computational Linguistics* 44(2). 197–235. doi 10.1162/COLI\_a\_00314
- Nivre, Joakim. 2005. Dependency grammar and dependency parsing. Relatório técnico. School of Mathematics and Systems Engineering, Växjö University. MSI Report 5133
- Nivre, Joakim. 2006. Dependency parsing. Em *Inductive Dependency Parsing*, 45–86. Springer Netherlands. doi 10.1007/1-4020-4889-0\_3
- Nivre, Joakim & Sylvain Kahane. 2004. Some notes on the generative capacity of dependency grammar. Em *Workshop on Recent Advances in Dependency Grammar*, 45–52. ↗
- Pagani, Luiz Arthur. 2024. Duas noções de dependência. *Revista da Associação Brasileira de Linguística (ABRALIN)* 22(2). 163–186. doi 10.25189/rabralin.v22i2.2167
- Pagano, Adriana Silvina, Marta Deysiane Alves Faria Sousa & Jorge Baptista. 2024. Gramática de dependências: fundamentos e perspectivas. *Revista da Associação Brasileira de Linguística (ABRALIN)* 22. 158–162. doi 10.25189/rabralin.v22i2.2284
- Robinson, Jane J. 1970. Dependency structures and transformational rules. *Language* 46(2). 259–285. doi 10.2307/412278
- Sgall, Petr, Eva Hajicová & Jarmila Panevová. 1986. *The meaning of the sentence in its semantic and pragmatic aspects*. Springer Science & Business Media
- Souza, Elvis de, Magali Sanches Duran, Maria das Graças Volpe Nunes, Gustavo Sampaio, Giovanna Pedrino Belasco & Thiago Alexandre Salgueiro Pardo. 2024. Automatic annotation of enhanced universal dependencies for Brazilian Portuguese. Em *Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL)*, doi 10.5753/stil.2024.245342
- Tapanainen, Pasi & Timo Järvinen. 1997. A non-projective dependency parser. Em *5<sup>th</sup> Conference on Applied Natural Language Processing*, 64–71. doi 10.3115/974557.974568

- Tesnière, Lucien. 2015 [1959]. *Elements of structural syntax*. John Benjamins Publishing Company. doi [10.1075/z.185](https://doi.org/10.1075/z.185)
- Yli-Jyrä, Anssi. 2005. Approximating dependency grammars through intersection of regular languages. Em *Finite-State Methods and Natural Language Processing (FSMNLP)*, 76–87. doi [10.1007/978-3-540-30500-2\\_26](https://doi.org/10.1007/978-3-540-30500-2_26)
- Zwicky, Arnold M. 2006 [1985]. 13 Heads, Bases and Functors. Em *Heads in Grammatical Theory*, 292–316. Cambridge University Press