

O que é Sintaxe e por que ela é tão difícil?

João Paulo Lazzarini Cyrino

03/03/2021

Objetivos

- Entender a história da gramática no ocidente e a razão pela qual as coisas que se aprende em sintaxe na escola são como são.
- Explorar o que é sintaxe do ponto de vista linguístico.
- Compreender as dimensões de Análise, Descrição e Teoria da sintaxe enquanto uma disciplina linguística.
- Aprender o que está envolvido na prática de análise sintática e por que esse tipo de atividade pode ser tão difícil.

A gramática no ocidente

Gramática é uma palavra que tem vários significados a depender do contexto em que ela se encontra. Por exemplo, na escola, gramática é o estudo das regras que fazem o **português correto**. Já na linguística, gramática compreendido é o estudo das estruturas das línguas. Há definições mais amplas, que associam gramática a qualquer sistema que envolva regras para a formação de estruturas quaisquer: gramática das cores, gramática musical, gramática de gráficos, etc.

Fato é que existe uma história por trás de como concebemos, **no ocidente**, a língua como algo que segue, ou deve seguir, um sistema de regras. Enfatizo que isso se dá **no ocidente** pois existem outras tradições de compreensão das línguas (algumas que, infelizmente, não chegaram até nós) e é importante ter essa perspectiva. Antes de falar sobre essa história ocidental, vamos mencionar duas tradições que temos algum conhecimento:

- **Gramáticos indianos:** na Índia houve um período de preocupação com a sistematização do Sânscrito quando as narrativas Hindus estavam passando da oralidade para a escrita. Nesse ambiente há gramáticos bastante normativos, como Yaska (séc VI a.C.). Por outro lado, há Panini (séc V a.C.), que buscava descrever como o Sânscrito era falado, inclusive em suas diferentes variedades dialetais. Cada gramático seguia uma metodologia, mas o trabalho de Panini é bastante emblemático e utilizado em linguística até os dias de hoje.
- **Gramáticos chineses:** a tradição indiana se assemelha um pouco com a tradição ocidental sobre a qual já vamos falar. Na China, no entanto, as coisas se desenvolveram de forma bastante distinta. Houve um foco muito maior na descrição da pronúncia, provavelmente por conta da grande variação linguística na região e de os caracteres serem ideográficos. Interessante mencionar que o conceito de palavra foi exportado do ocidente para a China no século XVIII, o que trouxe aos chineses outras ideias sobre gramática.

Nosso vocabulário e forma de pensamento sobre as línguas deriva dos gregos. A gramática surge no contexto das escolas gregas, mais especificamente nos manuais de retórica. Tais manuais buscavam analisar os discursos

e mostrar como eles poderiam ser mais eficientes. Disso também se iniciou uma preocupação com a análise das próprias frases na língua, em busca de construções mais interessantes retoricamente.

Dessa época é de especial destaque a gramática de **Dionísio o Trácio** - *tékhne grammatiké* (Arte Gramatical) -, séc I a.C., de onde retiramos boa parte da terminologia gramatical que utilizamos até hoje. Importante mencionar que essa terminologia estava muito atrelada a características da língua grega. Mais tarde, os Romanos vieram a importar essa tradição e colaram essa terminologia para a língua latina (com algumas incompatibilidades). As gramáticas latinas tinham por objetivo padronizar o latim, de forma que as pessoas aprendessem a escrever o latim clássico e não o vulgar. Lembremos que o Império Romano era muito grande e registra-se que o latim tinha grande variação desde o início da expansão do território romano.

É importante considerarmos duas características dessas gramáticas fundacionais da tradição ocidental de analisar línguas:

- Forte ligação com a língua escrita. Elas estão focadas no texto escrito, fundamentalmente, que poderá vir a ser lido em um discurso, mas que será registrado de forma escrita.
- Forte ligação com a retórica. A gramática está a serviço da construção de frases que venham a ser retoricamente mais interessantes.

Essas duas características **sempre** estão presentes de alguma forma quando estamos fazendo análise gramatical, infelizmente. Do ponto de vista científico, elas introduzem um viés na análise gramatical e, infelizmente novamente, muitos linguistas não se dão conta disso. Pelo contrário, dada a origem de muitos linguistas, essas características, somadas ao que contaremos a seguir, são - de certa forma - até apreciadas.

Ao longo da Idade Média havia o ensino de artes liberais, sete disciplinas divididas em dois grupos: *Trivium* e *Quadrivium*. De especial interesse para nós é o *Trivium* que reúne *lógica, gramática e retórica*. Mais uma vez, a gramática está associada à retórica e, adicionalmente, à lógica. O *quadrivium* envolvia aritmética, música, geometria e astronomia. O Trivium era visto como o conjunto de disciplinas para o desenvolvimento das capacidades linguísticas. A gramática, no caso, era a gramática latina de Prisciano, bastante sistemática.

Ao final da Idade Média, gramáticas de línguas vernáculas começaram a ser escritas, fortalecendo as línguas dos reinos nacionais que estavam surgindo (Portugal, Espanha, França e Inglaterra). Essas gramáticas eram de caráter mais descritivo e muito interessantes de serem investigadas. Ainda assim, a educação era fortemente influenciada pelo Trivium e o latim era a língua em que as coisas deveriam ser expressas corretamente.

As coisas ficam mais interessantes após o influente racionalismo de Descartes. Em 1660 é publicada a *Grammaire générale et raisonnée contenant les fondamens de l'art de parler, expliqués d'une manière claire et naturelle*, por Antoine Arnaud e Claude Lancelot, do mosteiro de *Port-Royal-des-Champs*. Essa gramática é conhecida também por *Gramática de Port-Royal*. Enquanto o Trivium separava lógica, gramática e retórica, aqui a gramática vista como a expressão do pensamento, que deveria ser lógico no sentido cartesiano: deduções claras e corretas. Essa gramática trouxe uma série de novas definições para a terminologia anteriormente utilizada e atrelada às línguas latina e gregas. Essas definições tornariam a terminologia gramatical universal, já que a gramática se trata de uma expressão do pensamento.

A *Gramática de Port-Royal* consolida nossa concepção mais corriqueira sobre gramáticas:

- Falar bem é sinal de pensar bem
- A terminologia gramatical deve se aplicar a todas as línguas (é universal)

Ao longo dos séculos seguintes, esse universalismo foi responsável por descrições um tanto quanto estapafúrdias de línguas não-europeias, feitas com o propósito de traduzir a bíblia ou catequizar povos indígenas.

A **Sintaxe** tem um papel fundamental nessa concepção de gramática. A sintaxe, o ordenamento dos elementos da frase, é o que garante o fluxo do raciocínio. Nesse sentido, muitas das definições de sintaxe estão ligadas à lógica de Aristóteles. Um exemplo claro disso é a divisão da frase em *sujeito* e *predicado*, onde o primeiro é aquilo sobre o qual o segundo fala. Em termos mais lógicos: o *predicado* é a expressão que pode

ser a *verdade* do *sujeito*. Por exemplo, a expressão *é feliz* é a verdade de qualquer *sujeito* conectado a ela: *João é feliz* é verdade se *João* é feliz.

Por conta desse histórico, os conceitos utilizados hoje na sintaxe das escolas apresentam definições bastante complicadas, que estão ora associados ao racionalismo, ora à retórica, ora a alguma característica da língua grega ou latina que não existe no português.

Tudo poderia ser melhor com o Estruturalismo linguístico (início do século XX). O estruturalismo aboliu as concepções universalistas e racionalistas sobre a linguagem, favorecendo uma abordagem mais relacionada à língua tal qual é falada e ela em si como um sistema particular. Dessa forma cada língua se define pela sua própria estrutura. O Estruturalismo foi um período de grande avanço no que diz respeito ao entendimento que temos das línguas. Foi possível descobrir, com as descrições das diferentes línguas que eram feitas, uma imensidão de variedades sobre como o ser humano poderia comunicar certas ideias e o quão as línguas são diferentes entre si.

Entretanto, o Estruturalismo foi perdendo terreno a partir da metade do século XX, com os trabalhos de Chomsky (Gramática Gerativa) e Greenberg (Tipologia Linguística). O primeiro retomava um universalismo linguístico a partir de uma ideia de Gramática Universal, presente no cérebro dos seres humanos. Essa ideia teve origem no desenvolvimento da matemática e dos sistemas formais no século XX ganhou força no debate Chomsky x Skinner (behaviorismo). O segundo era uma crítica ao relativismo antropológico, que pregava que a variação das línguas e comportamentos humanos não teria limites. Greenberg mostrou, com uma metodologia específica, que é possível detectar esses limites e que há características universais na linguagem.

Hoje a linguística se encontra em um momento de transição. Aparentemente vamos voltar a uma visão mais particularista das línguas, com concepções mais interessantes sobre o universalismo. De qualquer forma, nas escolas **predomina a gramática de port-royal** e, mesmo que se fale mais e mais sobre preconceito linguístico, **os conceitos da gramática são preconceituosos** pois atrelam estruturas linguísticas ao **raciocínio correto**.

O que é sintaxe do ponto de vista da linguística?

Nós nos comunicamos por pronunciar uma sequência de sons ou sinais. Essa sequência de sons ou sinais segue ordenamentos tais que podemos interpretar significados a partir dela. Em linguística tentamos, justamente, explorar como que essas sequências de sons ou sinais que utilizamos se estruturam de forma a significar o que significam.

Podemos entender essas sequências em diferentes níveis (níveis de análise linguísticos). Em um primeiro nível, podemos estudar os sons em si: quais sons há na língua, quais variações de sons há, se há restrições de sons em determinados contextos, etc. Trata-se do nível fonológico. Em um segundo nível, podemos estudar combinações de sons mínimas às quais é possível atribuir algum significado. Essas combinações são denominadas morfemas e são estudadas no nível morfológico. Acredita-se - embora isso não seja mais aceito tão facilmente atualmente - que existem combinações de morfemas bastante sistemáticas nas línguas às quais damos o nome de palavras. Cabe à morfologia estudar essas combinações. A sintaxe, o terceiro nível, estuda as combinações dessas palavras de maneira que elas formem um enunciado linguístico, com uma interpretação proposicional (ou um ato de fala).

A sintaxe possui grande interface com a morfologia, muitas vezes tornando-se indistintas, e com a semântica (o quarto nível). A morfologia fornece a unidade básica da sintaxe: a palavra. A semântica fornece o ponto final da combinação de palavras: a proposição. As palavras, portanto, se combinam para formar frases (conjuntos de uma ou mais orações), que são interpretadas como proposições semânticas.

Ao longo das próximas aulas isso tudo ficará mais claro. Por ora é importante entender que neste curso estaremos focados no tipo mais básico de prática que envolve a sintaxe: a **análise sintática**. Ou seja, estaremos focados em entender como pegamos uma frase e dividimos ela em partes menores que contenham significados menores até chegar na palavra e, algumas vezes, veremos um pouco dentro da palavra também.

Sintaxe enquanto disciplina na linguística compreende também outros fazeres, como a **descrição sintática** e a **teoria sintática**. Esses são fazeres da pesquisa em Sintaxe. Para o profissional de Letras, no entanto, é

esperado que se entenda sob quais princípios podemos analisar as frases de uma língua e, dado um conjunto de conceitos sintáticos, ou melhor, dada uma gramática, como analisar uma frase segundo ela.

Análise, Descrição e Teoria Sintática

Sintaxe envolve, pelo menos, essas três dimensões de prática:

- **Análise Sintática:** dividir as frases em partes menores, dando nomes a essas partes de acordo com o que significam.
- **Descrição Sintática:** descrever as sistematicidades na formação das frases de uma determinada língua, ou de tipos de frases comuns a diferentes línguas.
- **Teoria Sintática:** estabelecer critérios e conceitos sobre como as frases da língua devem ser analisadas ou criadas, ou seja, criar gramáticas.

A *Análise Sintática*, na qual nos deteremos neste curso, depende de uma teoria que forneça os critérios que se deve utilizar. Na escola, essa teoria é normalmente a *gramática normativa*. A *gramática normativa* é uma péssima teoria científica na medida em que seus conceitos são definidos de maneira confusa (Gramática de Port-Royal) e ela está associada a uma língua artificial, um português padronizado. Há diversas teorias científicas sobre Sintaxe e veremos algumas delas nas próximas aulas, concretamente, a gramática gerativa, a gramática funcional-estrutural e a gramática estocástica.

De qualquer forma, a *análise sintática* pode ser compreendida de maneira mais fácil e independente de teoria, se entendermos alguns assuntos pressupostos, que sistematizo abaixo:

- Teoria dos Conjuntos e funções
- Lógica Formal de Predicados
- Morfologia (linguística)

Isso porque, na análise estamos pegando palavras (Morfologia) e incluindo-as em diferentes conjuntos (Teorias dos Conjuntos) que terão diferentes papéis (Funções) na formação do significado proposicional da frase (Lógica Formal). Se entendermos isso, veremos que boa parte das teorias (incluindo a gramática normativa) definem a análise nesses termos e tudo o que faremos é dar nomes para os diferentes conjuntos que formamos.

Por que Sintaxe é difícil?

Pode-se dizer que sintaxe é difícil em diferentes níveis. De qualquer forma, é importante sempre considerar que é natural sentir dificuldade nessa disciplina. Elenco duas dificuldades fundamentais:

A primeira dificuldade da sintaxe se deve ao fato de que, ao ser ensinada, é pressuposto que o aluno conheça a lógica de predicados e a teoria dos conjuntos que fundamentam o processo de análise. Muitas vezes nem o professor conhece isso. Dessa forma, a aula de sintaxe se torna um exercício de decorar nomes para padrões mal definidos de combinações de palavras e nada parece fazer muito sentido. É importante que o aluno sempre faça o auto-diagnóstico: o quanto estou tendo que *decorar* para aprender essa disciplina. Sintaxe não deveria exigir memória, mas sim raciocínio analítico.

A segunda dificuldade está relacionada à terminologia. Termos como sujeito e predicado, conforme vimos, estão enraizados na lógica. Já termos como oração e voz estão relacionados à gramática latina e grega e são bem mais fáceis de entender dado o funcionamento dessas línguas. Verbos, adjetivos, substantivos são definidos via características morfológicas. Adjuntos são especialmente difíceis de entender e possuem definições confusas como *termo não-obrigatório*.

Neste curso mitigaremos essas duas dificuldades adotando critérios *formais* (explicitamente baseados na lógica e na teoria dos conjuntos) de análise das frases e tentaremos traduzir os termos utilizados na gramática

normativa para esses critérios (na medida do possível). É importante entender que, ainda assim há toda a dificuldade de se desenvolver o raciocínio analítico que, infelizmente, é pouco estimulado na educação brasileira em detrimento de *decorar* para a prova.

A expectativa é que, após a trajetória neste curso, um tanto desafiadora mas também com seu lado divertido, o aluno ganhe segurança para lidar com análise sintática e facilidade em entender o que as gramáticas querem dizer. Isso abrirá caminho para uma maior facilidade no ensino de línguas (português e mesmo línguas estrangeiras) e também para aqueles que quiserem seguir adiante na pesquisa linguística.