2 - Gerativismo linguístico

Cid Ivan da Costa Carvalho

SciELO Books / SciELO Livros / SciELO Libros

CARVALHO, C. I. C. Gerativismo linguístico. In: CARVALHO, C. I. C., and BARBOSA, J. R. A., ed. *Teorias linguísticas: orientações para a pesquisa* [online]. Mossoró: EdUFERSA, 2021, pp. 43-69. ISBN: 978-65-87108-25-4.

<https://doi.org/10.7476/9786587108629.0004>.

All the contents of this work, except where otherwise noted, is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença [Creative Commons Atribuição 4.0](https://creativecommons.org/licenses/by/4.0/).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia [Creative Commons Reconocimiento 4.0](https://creativecommons.org/licenses/by/4.0/).

2

GERATIVISMO LINGUÍSTICO

Cid Ivan da Costa Carvalho

2.1 Considerações iniciais

Se a linguagem é vista como um traço essencial de nossa humanidade, então, parece razoável dizer que a linguística-ciência da linguagem-tem um longo caminho a percorrer, pois o seu objeto de estudo se mostra tão complexo que mais de 2.500 anos de pesquisa não foram capazes de depreender uma resposta clara sobre o que realmente é a linguagem. No entanto, ao longo desses anos, vemos alguns lampejos que nos auxiliam na compreensão desse objeto de forma parcial. Dentro do programa biolinguístico, acreditamos que a teoria gerativa lampeja um caminho de investigação muito mais abrangente sobre a linguagem e sobre os tópicos a ela relacionados.

Essa afirmação parece coerente, pois, se considerarmos inicialmente as seguintes indagações: Por que apenas nós humanos somos capazes de formar comunidades que vão além da noção de família? Por que somente nós, e nenhum outro animal, buscamos explicações para os eventos naturais, sociais, culturais e históricos? Por que só nós organizamos e planejamos o futuro de maneira que vai muito além do que as outras criaturas fazem? Por que exclusivamente os humanos produzem e apreciam ficções e fantasias e especulam, desenham e empregam formas artísticas às quais não são praticadas pelos outros animais? Parece que a única resposta plausível para essas indagações seja o fato de o homem possuir o fator mais importante que explica como essas capacidades notáveis chegaram até nós—a linguagem.

McGilvray (2014) diz que essa conclusão faz sentido porque a linguagem fornece a capacidade de especular e pensar sobre algum conteúdo próximo ou distante, construindo um número ilimitado de pensamentos complexos, passíveis de separar das circunstâncias e estender sobre tempos e circunstâncias arbitrárias. Ela fornece a capacidade de organizar e construir projetos que envolvem a cooperação entre os indivíduos. Porque, sem ela, não se pode fazer ciência nem buscar explicações básicas e conceitos matemáticos. Essas e outras capacidades estão diretamente relacionadas ao fato de sermos possuidores da linguagem.

Se considerarmos também que a linguística é uma ciência natural objetiva que trata a linguagem como um sistema biológico que evoluiu a partir de um único indivíduo e foi geneticamente transmitido a sua prole, diríamos, então, que as línguas não são ferramentas que os seres humanos projetam ao longo dos milhares de anos, mas são objetos biológicos, como o sistema virtual ou imunológico digestivo, por exemplo, como diz Chomsky (2018a).

Considerando esses pontos fundamentais para o estudo da linguagem humana, percebemos que a linguagem é mais que um sistema de comunicação ou de interação; ela é mais que um depósito de imagens de

palavras na mente dos membros de uma comunidade; ela é mais que um contrato estabelecido entre os membros de uma comunidade de fala; ela é mais que uma série de hábitos usados para responder a situações de fala convencionais e para responder aos sons com as ações; ela é mais que um o corpo de sinais proferidos e audíveis pelos quais o pensamento é expresso; ela é mais que um método puramente humano e não instintivo de comunicar ideia e emoções por meio de símbolos produzidos voluntariamente.

Nesse sentido, o primeiro tópico deste capítulo apresenta o surgimento da teoria gerativa e sua importância na década da revolução cognitiva. O segundo tópico mostra os principais conceitos que julgamos necessários à compreensão geral da teoria. Todavia, os leitores que pretendem se aprofundar no assunto encontrarão nas referências as principais obras relacionadas ao tema. O terceiro tópico expõe a metodologia empregada na linguística gerativa, fazendo a distinção entre o método utilizado pelos gerativistas e os métodos aplicados pelos distribucionalistas. O quarto tópico apresenta um recorte de uma pesquisa com a utilização da teoria gerativa.

2.2 Surgimento da teoria

A revolução cognitiva foi um movimento intelectual que se iniciou na década de 1950, como proposta interdisciplinar de estudar a mente e os processos a ela relacionados. As áreas mais relevantes eram a psicologia, a linguística, a ciência da computação, a antropologia, a filosofia e a neurociência, nas quais predominavam o pensamento behaviorista. O behaviorismo entendia que a psicologia humana podia ser estudada com base na observação do comportamento individual. Para isso, o comportamento era testado em laboratório. Nesse sentido, essa perspectiva tomava o comportamento como um dado que evidenciaria os mecanismos internos da mente e esses mecanismos operariam ao executar as ações e ao interpretar a experiência. A revolução cognitiva surgiu em oposição a essa concepção e as principais publicações que desencadearam o movimento foram a obra

The Magical Number Seven, Plus or Minus Two (1956), do psicólogo George Miller, e a obra *Syntactic Structures* (1957), do linguista Noam Chomsky.

Nessa época, os estudos estruturalistas do linguista Leonard Bloomfield constituam a baliza que definia a prática científica da linguística. A linguagem humana era vista como um dos fatores sociais condicionados e sua aquisição dependia dos estímulos e respostas dos sujeitos na interação social. Assim, considerando os pressupostos da teoria behaviorista, uma comunidade ensina o indivíduo a emitir uma dada resposta verbal (uma palavra), provendo estímulos reforçadores na presença da coisa para a qual o termo proferido é tomado como referente. Nesse sentido, Bloomfield (1935) aponta que a aquisição da linguagem é construída por meio dos eventos práticos que precedem o ato de fala; depois, ocorre a produção da própria fala e, por último, há os eventos práticos seguidos de ato de fala. Podemos simplificar com a seguinte sequência: temos os estímulos e as respostas aos estímulos, que serão estímulos para outras respostas.

Pensando assim, Costa (2016, p. 123) afirma que o distribucionalismo, como era chamado o estruturalismo norte-americano, pressupunha: que cada língua apresentava uma estrutura específica, pois fazia parte de uma comunidade de fala; que os níveis fonológico, morfológico e sintático evidenciavam essa estrutura; que as unidades de nível imediatamente inferior constituem a unidade seguinte; que a descrição de uma língua deve começar pelas unidades mais simples; que a posição estrutural linear da unidade linguística define a função da unidade; que, na descrição, é necessária absoluta objetividade; que o processo de combinação das unidades é guiada por leis próprias do sistema linguístico. Esses pressupostos auxiliavam os pesquisadores distribucionalistas no estudo das línguas até meados dos anos de 1950.

No entanto, em *Estrutura Sintática* [*Syntactic structures*], Noam Chomsky lança as bases da teoria gerativa como uma proposta diferente os distribucionalistas americanos, tanto na concepção de língua natural quanto nos pressupostos teóricos defendidos por eles. Essa teoria procura

descrever as línguas particulares e explicar abstratamente o funcionamento das formas linguísticas da linguagem.

Ao contrário dos distribucionalistas, Chomsky (2015, p. 17) afirma que “[...] uma língua é um conjunto (finito ou infinito) de sentenças, cada sentença sendo finita em extensão e construída a partir de um conjunto finito de elementos”. Ele acrescenta que todas as línguas naturais são línguas nesse sentido, uma vez que cada língua apresenta um número finito de fonemas e as sentenças podem ser definidas como um conjunto desses fonemas. Essa afirmação corrobora com o fato de que as línguas apresentam uma estrutura comum entre si. Essa estrutura comum é evidenciada a partir da estrutura sintática, pois é na sintaxe que se encontram os princípios e processos que constituem as sentenças das línguas particulares. Um nível linguístico, como o fonológico ou o morfológico, é essencialmente um conjunto de recursos descritivos que estão disponíveis para a construção de uma gramática. Explicando o objeto de estudo proposto por Chomsky, Barbosa (2013, p. 21) comunica que, nessa concepção, “caberia ao linguista descrever as línguas naturais com vistas a determinar as sequências universais da língua, comuns a toda e qualquer língua”.

Por esse motivo, a descrição de uma língua deve começar pela sintaxe, unidade mais abrangente no nível de análise, pois ela determina a construção de uma gramática para língua particular, ou seja, o mecanismo de produção de sentenças de determinada língua. No entanto, como determinar a gramaticalidade de uma sentença? Para Chomsky (2015), o objetivo básico da análise linguística de uma língua qualquer é separar as sentenças gramaticais das sentenças agramaticais de uma língua e estudar a estrutura das sentenças gramaticais.

Além disso, cada unidade é definida na forma estrutural hierárquica e não na estrutura linear das sentenças. Vejamos as sentenças abaixo:

- (1) a) O homem chutou a bola de plástico.
- b) O homem chutou a bola de tarde.

Podemos observar que as sentenças 1a e 1b pertencem à gramática da língua portuguesa e que os sintagmas preposicionais (*prepositional phrase*–PP) “de plástico” e “de tarde” se encontram na posição linear da estrutura nas sentenças, ou seja, antes do nome (*noun*–N) “bola”. Todavia, o sintagma preposicional “de tarde” está numa posição acima, na estrutura hierárquica, uma vez que ele não pertence ao sintagma nominal (*noun phrase*–NP) encabeçado pelo item lexical “bola”, mas ao sintagma verbal (*verbal phrase*–VP) encabeçado pelo verbo “chutar”, como ilustra a figura 1.

Figura 1 – Ilustração arbórea das sentenças 1a e 1b

Fonte: Sistema Constituency Parser do Lx Center (2020).

Outro posicionamento do gerativismo, cujo cerne já se encontrava nas obras iniciais de Chomsky, é o fato de que o processo de combinação das unidades é guiado por leis universais, desconsiderando as particularidades das línguas, pois “a linguagem é um objeto natural, um componente da mente humana, representado fisicamente no cérebro e integrado ao patrimônio biológico da espécie. (CHOMSKY, 2018b, p. 1).

Para a perspectiva gerativa, o que significa dizer que a linguagem é um objeto natural, como Chomsky a definiu? Que propriedades da linguagem

são estudadas quando a consideramos um objeto natural? Que concepção de linguagem é adotada na teoria gerativa? Que conceitos estão presentes nessa teoria? No próximo tópico, apresentaremos algumas explicações que podem levar às respostas dessas perguntas.

2.3 Princípios teóricos

Quando Chomsky (2018b) diz que a linguagem é um objeto natural, ele pressupõe que “se os humanos são parte do mundo natural e não seres sobrenaturais, então a inteligência humana tem seu escopo e seus limites determinados pelo design inicial.” (CHOMSKY, 2005, p. 195). Esse é um ponto fundamental na teoria gerativa, pois esta concebe inicialmente que o homem é um ser biológico e os aspectos mentais, como linguagem, pensamento, percepção, memória, etc. são objetos naturais comparados a moléculas, campo elétrico, sistema respiratório. Não podemos fazer distinção entre a linguagem e as moléculas de H₂O, por exemplo, pois ambos os objetos são naturais. Nesse sentido, não se pensa que o primeiro é cultural e o segundo é natural. Os hábitos culturais são manifestações exteriores que a própria natureza do ser pode expressar. Quando olhamos um pássaro, por exemplo, um João-de-barro, vemos que essa ave é um ser natural, como outros objetos também os são. Se observarmos o ninho que eles fazem nos galhos das árvores, percebemos que o material utilizado para a construção é feito de material natural (barro e gravetos), porém não diríamos que o ninho é um objeto natural. Se escutarmos o canto desse pássaro, diríamos que o canto seria um objeto natural ou cultural? Ele aprende o canto porque escuta os outros da mesma espécie cantarem ou ele tem a competência adquirida da espécie para desempenhar o canto? O gerativismo entende que os objetos naturais são distintos dos objetos culturais. O ninho de um pássaro é um objeto “cultural”, pois é construído por meio de um processo de transformação da matéria prima, mesmo

que esse ninho seja uma escavação em rocha ou em terra. Todavia, não diríamos o mesmo da capacidade de construí-lo.

Quando recorremos à linguagem numa perspectiva naturalista, diremos que ela é uma faculdade, um dos elementos da mente/cérebro dos seres humanos, uma estrutura especializada da mente/cérebro com a função de executar os processos linguísticos. A faculdade da linguagem é o estado inicial geneticamente determinado para todos os indivíduos: para Peter, Juan, John e outros seres humanos. Esse estado Inicial "determina de maneira singular o sistema computacional da linguagem, junto com uma série muito estruturada de possibilidades lexicais e algumas opções entre 'elementos gramaticais' que não apresentam conteúdo substantivo. (CHOMSKY, 2005, p. 66).

A faculdade da linguagem é um módulo da mente especificamente associado à língua, e não a outras linguagens (como a pintura, a música, a dança, etc.) e apresenta, no mínimo, dois componentes diferentes: “um ‘sistema cognitivo’, que armazena as informações de alguma maneira, e sistemas de desempenho, que fazem uso dessas informações para articulação e percepção, para falar sobre o mundo, fazer perguntas, contar piadas e assim por diante” (CHOMSKY, 2005, p. 210). Define-se o primeiro componente como sendo: um corpo comum de informações inatas que os sistemas de desempenho acessam, ligam-nos e os provêm de instruções; a parte da linguagem que permanece intacta; a mais básica propriedade da linguagem; uma dotação genética, chamada de estado inicial, que torna possível o desenvolvimento de qualquer língua, a Gramática Universal. O segundo componente é descrito como sendo: uma parte da linguagem que é mutável e “seletivamente imperfeita”; uma unidade totalmente determinada pelo estado inicial da linguagem; um procedimento computacional que satisfaz a propriedade básica. Destacamos que a expressão estado inicial se refere ao estado em que a linguagem está quando a criança nasce, sendo considerada uniforme em relação a toda a espécie humana.

Observando bem a distinção desses dois componentes da linguagem, vemos que tais componentes estão diretamente associados à Língua-I (interior, individual, intensional) e a Língua-E (exterior, social, extensional). A língua-I é um estado do sistema cognitivo da linguagem de um indivíduo; semelhante a maneira de falar de uma pessoa. A língua-I, por exemplo, de uma criança nos estágios iniciais de aquisição apresenta forma de falar diferente dos adultos. Se a língua-I fosse um objeto “palpável” e abrissemos para ver de que ela é composta, encontraríamos três sistemas integrados: o sistema computacional, o sistema sensorio-motor e o sistema conceitual-intencional, como descrita por Chomsky e Berwick (2017, p. 20):

- (1) um sistema computacional interno que constrói expressões hierarquicamente estruturadas com interpretações sistemáticas no nível das interfaces com outros dois sistemas internos, a saber: (2) um sistema sensorio-motor para a externalização como produção ou *parsing* e (3) um sistema conceitual para inferência, interpretação, planejamento e organização de ação—o que formalmente é chamado de ‘pensamento’.

O sistema computacional interno mapeia um leque de itens lexicais em um par de objetos simbólicos que apresentam as características fonéticas (som) e os aspectos semânticos (significado). Considerando que S corresponde a um significado no português e que F se refere a uma cadeia de sons, S e F fornecem as informações necessárias para os sistemas de desempenho. Veremos, logo a seguir, o que são esses sistemas. Por ora, diremos que o sistema computacional executa esse mapeamento por meio de uma operação chamada de *Merge*.

A operação das expressões hierarquicamente estruturadas na Língua-I é feita por meio de *Merge*. Esse é o princípio computacional básico de todas as línguas. Conforme a ideia exposta por Chomsky (2014, p. 34 - 35), há dois tipos de *Merge* possíveis: o *Merge* que toma “duas coisas e colocá-las juntas” sem que uma altere o estado da outra. E o outro *Merge*, que toma

“uma coisa e um pedaço dela” e prende “esse pedaço na borda ou margem [edge] da coisa”. Essa é a propriedade de deslocamento ou movimento das línguas naturais, que é encontrado em todos os lugares”. O primeiro tipo é chamado de *Merge* interno e o segundo, de *Merge* externo. “*Merge* externo é usado, em essência, para fornecer as estruturas de argumentos. *Merge* interno é basicamente usado para fornecer informação relacionada ao discurso como foco, tópico, informação nova—todo tipo de coisa que se relaciona com a situação do discurso.” Nesse sentido, *Merge* é uma operação computacional que toma dois objetos humanos compostos de som e significado, como ($x = \text{ler}$) e ($y = \text{livro}$), e os concatena para formar o terceiro objeto ($z = \text{ler livro}$) sem alterar a estrutura de nenhum desses objetos. Os objetos “ler”, “livro” e “ler livro” permanecem os mesmos. Essa operação é fundamental para a produção das expressões estruturadas sintaticamente, que são interpretadas no sistema conceitual-intencional e emitidas pelo sistema sensorio-motor.

Observe que *Merge* opera com os itens lexicais. Mas o que são itens lexicais na teoria gerativa? Jackendoff (2002) define o item lexical como uma unidade armazenada na memória de longo prazo do falante de extensão maior ou menor que as palavras gramaticais. Para esse autor, nem todas as palavras gramaticais são itens lexicais e a gramática universal pode ser formulada como uma coleção de itens lexicais abstratos que “semeiam” a aquisição da linguagem. Seguindo a mesma ideia de Jackendoff, Rio-Torto (2016) dispõe que a organização das estruturas dos itens lexicais tem lugar na mente do falante de modo dinâmico, uma vez ele analisa as estruturas antes de utilizá-los na formação de novas palavras. É claro que as palavras escritas são representações fixas do léxico do falante, constituindo parte do material linguístico produzido, quando expressa o pensamento com a utilização do léxico ou cada vez que ele emite uma produção da fala ou da escrita.

O falante pode fazer a montagem do lexema em cada ato de fala. Para isso, o falante recorre aos mecanismos paradigmáticos, às

bases e aos operadores afixais disponíveis para formar nomes deverbais de 'ação'. Escolhe a base que encerra o significado conceptual que pretende exprimir e para essa base escolhe o afixo que obedece aos constrangimentos que a base e o padrão derivacional impõem (RIO-TORTO, 2016, p. 81).

Além dessas características, Kenedy (2013) expressa que o léxico é o componente inicial do processamento computacional da linguagem que contém a forma fonética (π) e a forma lógica (λ), ou seja, no léxico, encontram-se armazenadas todas as informações relacionadas à emissão de som e ao significado, que devem ser mescladas no espaço derivacional. Ele tem a função de fornecer à sintaxe os itens lexicais com os quais o sistema computacional gera as sentenças da língua.

Resumindo o que foi dito acima, o léxico, numa perspectiva cognitiva da linguagem, possui instâncias de interface com a memória de longo prazo, com os sistemas conceitual e articulatório e com o sistema computacional; é nele que a língua-I e os demais sistemas cognitivos e as contingências da língua-E se entrecruzam. Como menciona Kenedy, (2013), o léxico existe na cognição individual e reflete as relações conceituais e morfossintáticas, derivando de convenções socioculturais e históricas. Os valores conceituais de uma língua são codificados no léxico de seus falantes.

Nesse sentido, a faculdade da linguagem assume a forma de uma Língua-I completa e se integra num sistema de desempenho que executa um papel na articulação, na interpretação e na expressão de crenças e desejos, referindo, contando histórias e assim por diante. (CHOMSKY, 2005). Ela assume essa forma, uma vez que cada falante a utiliza no desempenho da língua-I.

Os sistemas de desempenho estão associados à Língua-E (exterior, extensional) e se apresentam em dois tipos distintos: o sistema sensorio-motor e o sistema conceitual-intencional. Uma expressão linguística é gerada por meio desses dois níveis de interfaces: "uma que fornece informação e instrução para os sistemas articulatório-perceptivos e outra para os sistemas

conceitual-intencionais". (CHOMSKY, 2005, p. 67). O primeiro fornece as instruções para a externalização da língua que são caracterizadas como a representação fonética (forma fonética–FF), ou seja, é responsável pelo controle da produção e da recepção dos sons na fala e dos sinais visuais nas línguas vísuo-espaciais. O segundo fornece informações para os processos mentais e são caracterizados como a representação lógica (forma lógica–FL), ou seja, é responsável pelas crenças, desejos, conceitos, lógica.

Se observarmos os conceitos de Língua-I e de Língua-E com mais detalhes, perguntaremos: o que há de comum entre um falante do português e um falante do mandarim? Diríamos que é o fato de ambos compartilharem a única estrutura virtual dos seres humanos, que é o estado inicial da faculdade da linguagem. No entanto, o que os diferencia são os sistemas de desempenho que eles adquiriram e aplicariam a casos específicos no uso da língua-E.

Parafraseando Chomsky (2005), quando perscrutamos as propriedades fonéticas das Línguas-E, encontramos uma textura rica de elementos sonoros puramente internalista e evidências para as relações fonéticas entre as línguas. Uma grande parte dessa estrutura fonética deriva de nossa natureza interior, determinada pelo estado inicial de nossa faculdade de linguagem, por isso, não apreendida e universal para as Línguas-I. Como diz Chomsky (2005, p.79), "Algumas propriedades da expressão são universais, outras são particulares da linguagem." Nesse sentido, as propriedades fonéticas do português, no nível da interface articulatorio-perceptual, apresentam muitas semelhanças com outras línguas, porque, dos sistemas de desempenho, a representação no nível da interface na forma fonética impõe condições estruturais sobre o uso, ou seja, sobre a articulação e a percepção.

Embora a capacidade de controlar o trato vocal para a fala pareça ser uma propriedade especificamente humana, isso não é considerado determinante na produção da linguagem, uma vez que a produção da linguagem humana não depende da modalidade de externalização,

o que as línguas de sinais deixam bem claro. A externalização, então, é um processo auxiliar e suas propriedades, um reflexo do sistema sensorio-motor. Conforme Chomsky (2018a, p. 43), “[...] o processamento é um aspecto periférico da linguagem e que usos específicos da linguagem que dependem da externalização, entre eles a comunicação, são ainda mais periféricos, contrariamente ao dogma virtual que não tem nenhum embasamento sério.”

Esses conceitos nos levam a compreender melhor o que a teoria gerativa demonstra sobre os princípios e os parâmetros da linguagem. Os princípios da linguagem são fixos e inatos ao ser humano e os parâmetros são as variações restritas aos sistemas de desempenho. Segundo Chomsky (2014), inicialmente, a teoria gerativa sustentava que um parâmetro era uma opção disponível em um princípio linguístico, uma “regra” universal. No entanto, nas discussões mais recentes da teoria, o parâmetro se apresenta como duas formas distintas, a saber: as categorias funcionais e os micro-parâmetros.

Um dos principais exemplos dos parâmetros de categorias funcionais é o parâmetro sobre a ordenação entre o núcleo, o especificador e os complementos que, geralmente, levam as línguas a escolherem um parâmetro de ordenação para as categorias de sujeito (S), verbo (V) e objeto (O). O português, o francês e o inglês escolhem, para essas categorias, que o núcleo do *Inflectional phrase* (IP) precede o complemento (Compl.) e o especificador (Spec.) precede o núcleo (V). Diferentemente dessas línguas, o alemão ordena essas categorias de forma inversa, de modo que o objeto (O) antecede o verbo (V) e o sujeito (S) antecede o objeto.

A figura 2 ilustra a diferença desses dois parâmetros.

Figura 2 – Ilustração do parâmetro da ordem nas categorias de sujeito, de verbo e de objeto no francês e no alemão, respectivamente

Fonte: Adaptação de Raposo (1993).

Para Raposo (1993), numa língua em que o valor do parâmetro é [X + Compl], como o francês, a ordem canônica é VO, como ilustra a figura em 2a; numa língua em que o valor do parâmetro é [Compl + X], como o alemão, a ordem canônica é OV, como ilustra a figura em 2b. Percebe-se aí que, quando o verbo fica em posição final, o complemento e o especificador aparecem à sua esquerda ou o objeto direto pode aparecer à sua direita. No entanto, quando o verbo fica no meio, o objeto direto pode aparecer à sua esquerda.

Essas distinções na ordem das categorias funcionais apontam para o fato de que as diferenças paramétricas entre as línguas são, na verdade, diferentes formas que a faculdade da linguagem tem disponíveis para preencher os *outputs* estabelecidos pelos sistemas de desempenho, pelas Línguas-E, com os quais ela “interage” mediante as interfaces sensório-motora e conceitual-intencional. Presume-se também que as diferenças paramétricas estejam alojadas nas palavras funcionais, ao invés de nos princípios em si mesmos, como era sugerido na abordagem inicial, conforme afirma Chomsky (2014).

Os micro-parâmetros identificam as unidades mínimas de variação paramétrica nos itens lexicais. Kayne (2005, p. 4) afirma que uma ideia generalizada sobre os parâmetros sintáticos foi a de que eles estão limitados a serem características/propriedades de elementos funcionais, em vez de serem características de elementos lexicais. Porém, como os elementos funcionais fazem parte do léxico, essa limitação significa dizer que os parâmetros sintáticos são necessariamente características ou propriedades dos elementos do léxico. Essa ideia apresentada por Kayne “veio a ser casada também com a ideia de que todas as diferenças paramétricas estão localizadas no ‘léxico’” (CHOMSKY, 2014, p. 402).

Como vimos, os parâmetros dizem respeito às propriedades específicas das línguas e os princípios são propriedades universais presentes nas línguas. Uma dessas propriedades no nível sintático é que os sintagmas X tem X como núcleo, ou seja, o NP (*Noun Phrase*) tem um N como seu núcleo; em outras palavras, cada categoria sintagmática é projetada por seu núcleo. Disso resulta o seguinte princípio: todo sintagma é a projeção de seu núcleo. Com base nesse princípio, dizemos que a estruturação hierárquica dos constituintes é basicamente a mesma para todos os sintagmas. Mais precisamente, os núcleos N, V, A ou P admitem modificadores e complementos que ocorrem na mesma ordem para todas as categorias. Em português, todos os modificadores ocorrem em princípio à esquerda do núcleo e os complementos, à direita do núcleo.

Trazendo a ideia dos princípios para o nível fonológico, dizemos que o número de segmentos que podem ser associados a determinado constituinte silábico e a ordem em que tais segmentos ocorrem são definidos pelas restrições dos princípios de sonoridade, da distância mínima e do máximo *onset*.

Para Silva (2014, p.207), o princípio da sonoridade é “uma gradação referente ao grau de abertura do trato vocal durante a produção dos sons e da quantidade de energia produzida durante a produção de um som”. Tal princípio expressa a tendência de que os segmentos mais sonoros ficam

mais próximos do núcleo silábico e os menos sonoros ocupam as posições periféricas. Em português, as líquidas são mais sonoras que as nasais e que as obstruintes e, por isso, elas estão mais próximas do núcleo num ataque complexo e na coda complexa.

A posição periférica é ocupada pelas consoantes líquidas, nasais e obstruintes. O português apresenta uma distância mínima menor que a líquida ou igual a ela para a segunda consoante do ataque complexo. Na sequência de segmentos /pl/, como na palavra “placa”, há um aumento de sonoridade que sai da obstruinte /p/ para a líquida /l/. Todavia, quando a realização da sequência é, por exemplo, /pn/, como na palavra “pneu”, em que a sonoridade aumenta de uma obstruinte para uma nasal, há a inserção de uma vogal entre as duas consoantes – [pinew] ou [penew]. Esse princípio explica a realização da epêntese e a formação do ataque e da coda complexos em português. (CARVALHO, 2016).

O princípio do máximo *onset* diz que uma sílaba espelha a coda no nível de sonoridade, pois enquanto o primeiro cresce em sonoridade, o segundo decresce, ou seja, o nível de sonoridade do onset vai das obstruintes às líquidas e o nível da coda vai das líquidas às obstruintes. A sílaba /təans/ na palavra “transporte”, por exemplo, exemplifica muito bem esse princípio, pois o nível de sonoridade aumenta da obstruinte /t/ para líquida /ə/ e depois para a vogal /a/ e diminui da vogal /a/ para a nasal /n/ e depois para a obstruinte /s/.

No próximo tópico, veremos os processos metodológicos que são utilizados na pesquisa gerativa, bem como as técnicas e os procedimentos aplicados pela teoria para obtenção de resultados.

2.4 Metodologia para a pesquisa

Os métodos de uma ciência se caracterizam pela constituição de um conjunto das etapas que o pesquisador deve seguir na realização de uma pesquisa. Os objetos de investigação determinam o tipo de método a ser

empregado. No entanto, as técnicas que compõe o método podem ser distintas. Quando pensamos a linguagem como o objeto de investigação do linguista, percebemos que existem diferentes aplicações dos métodos e das técnicas empregados por ele na realização da pesquisa. Apontaremos aqui o método na teoria distribucionalista e detalharemos o método na teoria gerativa.

Os pressupostos teóricos dos distribucionalistas, que apresentamos no início deste capítulo, guiavam os processos metodológicos da pesquisa na seguinte direção: (1) era necessária a constituição de um *corpus* mais variado possível de ocorrências linguísticas da comunidade de falantes da língua; (2) a partir do *corpus*, elaborava-se um inventário que permitisse identificar as unidades elementares de cada nível; (3) depois, verificavam-se as leis combinatórias dos elementos linguísticos em cada nível. Esse método de pesquisa é conhecido como método indutivo, no qual o linguista parte das observações e dos experimentos para fazer generalizações sobre os fatos linguísticos, ou seja, o linguista passa dos fatos particulares observados na língua para aplicar a todos os fatos da mesma espécie. O que se constata da amostra no *corpus*, ele estende a toda a população de casos da mesma espécie na língua.

A técnica de análise que eles utilizavam era chamada de análise de constituintes imediatos. Essa técnica consiste em decompor o *corpus* em unidades significativas, em frases; depois, a frase segmentada em constituintes, chamados de sintagmas; em seguida, o sintagma é dividido em palavras; por último, a palavra era separada em morfemas e fonemas. Nesse sentido, os distribucionalistas eram indutivistas na metodologia de investigação e, na prática científica, eram descritivistas.

No entanto, para compreendermos o processo metodológico da teoria gerativa, devemos olhar os objetivos dessa teoria. Como apresentamos acima, a noção de língua está associada ao estado inicial da faculdade da linguagem e aos resultados do desenvolvimento desse estado inicial, pelo contato com um determinado ambiente linguístico do falante. Nesse

sentido, a teoria gerativa tem dois objetivos fundamentais: (1) descrever o conhecimento do falante da língua em particular, por exemplo, o português (ou a língua de sinais brasileira); (2) explicar os processos mentais do falante ideal que sai do estágio inicial do conhecimento linguístico inato até o conhecimento pleno da língua.

Tendo em vista esses dois objetivos, veremos o método científico utilizado e as técnicas empregadas na teoria gerativa.

2.4.1 Método científico empregado no gerativismo

O gerativista utiliza o método dedutivo (hipotético-dedutivo). Esse método considera que a pesquisa científica é uma livre criação de hipótese e que a descoberta científica é resultado da construção da hipótese. Pensando assim, a pesquisa em linguística gerativa segue as seguintes etapas: (1) “formula-se uma hipótese, uma conjectura acerca do objeto, sob a forma de lei geral;” (2) “procura-se tirar conclusões por meio de deduções lógicas, ou seja, faz a dedução das consequências da lei;” depois, (3) compara-se “essas consequências entre si e com o que é observado, a fim de se verificar a adequação da hipótese proposta.” (LOBATO, 1986, p. 79). Essa autora afirma também que o gerativista e o distribucionalista se diferenciam pelo fato de o método dedutivo permitir a introdução de hipóteses na teoria e a aceitação do uso da intuição na análise científica. Por um lado, os estruturalistas americanos não aceitavam a construção das hipóteses nas pesquisas, pois, para eles, o linguista não deveria questionar os dados linguísticos, mas apenas descrevê-los sem fazer qualquer indagação. Por outro lado, com a utilização do método hipotético-dedutivo, a linguística propõe conjecturas sobre a linguagem e as línguas. Ao invés de limitar a descrição dos dados de um *corpus*, como faziam os distribucionalistas, o linguista gerativista propõe hipóteses as quais devem ser testadas e postas em confronto com os dados observados nas línguas. Essa característica da pesquisa gerativista privilegia dois aspectos atribuidores de cientificidade

da linguística: o primeiro é “o carácter explicativo (quando afirma que a linguística faz ‘predições’)” e o segundo é “o carácter empírico (quando diz que essas predições são ‘empiricamente testáveis’)” (LOBATO, 1986, p. 80).

Essas duas características estão diretamente relacionadas aos objetivos propostos pela teoria, como apresentamos acima. O gerativista utiliza a descrição das línguas, que constitui a habilidade de fazer com que o outro veja mentalmente aquilo que o pesquisador observou em campo. Essa descrição deve ser suficientemente precisa para que o leitor visualize o que o linguista observou em sua pesquisa. Nesse sentido, ela pode ser um passo metodológico utilizado pelo linguista e pode ser uma técnica da pesquisa, de modo que outro pesquisador possa replicá-la, aplicando o mesmo método. A demonstração das descrições linguísticas é feita por meio das árvores sintáticas ou dos diagramas arbóreos.

Além da descrição, o linguista gerativista procura explicar o conjunto de propriedades genéticas, inatas e universais presentes na gramática universal (GU), ou seja, procura explicar como há diversidade de línguas se somos dotados por um único aparato genético. A explicação é feita por meio das noções de princípios e parâmetros. A faculdade da linguagem é construída por leis gerais que servem para todas as línguas naturais. Um fenômeno que viola os princípios gerais não é tolerado em nenhuma língua particular. Talvez isso ocorra devido à forma como a mente/cérebro da espécie humana funciona. É tarefa do gerativista explicar como os processos, as estruturas linguísticas acontecem na mente do falante.

2.4.2 Técnicas científicas utilizadas no gerativismo

As técnicas de coleta de dados e os procedimentos operacionais que servem de mediação prática em pesquisa gerativa constituem elementos precípuos. Eles são instrumentos que auxiliam o pesquisador na provação ou não de sua hipótese. Abaixo, apresentaremos três principais técnicas utilizadas na pesquisa gerativa.

Mencionaremos, em primeiro lugar, que a base empírica para a construção da gramática de uma língua é a introspecção acerca das entidades linguísticas, ou seja, os juízos de gramaticalidade e aceitabilidade da língua feitos pelos pesquisadores. Essa afirmação pressupõe que o pesquisador é falante plenamente competente da língua e que a sua competência é idêntica à dos demais falantes.

Outra fonte muito importante é a documentação sobre a língua ou os *corpora* de textos. Os *corpora* são registros e a sistematizações dos dados, informações organizadas em condições de serem analisadas pelo pesquisador. No contexto de realização de uma pesquisa gerativa, a documentação é uma técnica de identificação, levantamento e exploração de documentos-fonte do objeto pesquisado e registro do uso pelo falante. As entidades linguísticas que se encontram nos *corpora* ajudam a sugerir informações sobre a língua que escapam ao escrutínio mental do pesquisador.

Além dessas, podemos mencionar a entrevista *in loco*. Com a aplicação dessa técnica, o linguista gerativista colhe as informações de um determinado fenômeno linguístico diretamente dos falantes da língua. Com a utilização dessa técnica, o linguista procura compreender o que os falantes pensam, sabem, representam sobre o conhecimento linguístico da língua que eles falam. Para isso, ele pode utilizar uma entrevista não diretiva ou a entrevista estruturada. Na primeira, o linguista colhe as informações dos falantes a partir do discurso livre, deixando-os muito à vontade para expressar seus conhecimentos sobre as representações linguísticas; na segunda, o linguista direciona as perguntas com questões diretivas, de modo a obter os resultados esperados pela pesquisa.

No próximo tópico, apresentaremos um recorte de pesquisa que auxilia o leitor na compreensão e na prática científica desta teoria.

2.5 Recorte de uma pesquisa

Vimos, no tópico anterior, que as hipóteses são conjecturas acerca do objeto, sob a forma de lei geral. Quando definimos a hipótese de nosso trabalho, temos de avaliá-la conforme a coerência interna, consistente das noções usadas e a utilidade das hipóteses na análise das estruturas da língua, como menciona Perini (2019). Antes de sugerirmos uma hipótese, vejamos uma sentença explicada por esse autor.

(2) Os passarinhos atacaram a cobra.

Observamos na frase (2) alguns aspectos formais: (a) o constituinte “os passarinhos” aparece em primeiro lugar, antes do verbo; (b) o constituinte “a cobra” aparece depois do verbo; (c) o verbo concorda em número e pessoa com o constituinte que o antecede. Esses elementos definem as funções sintáticas dos sintagmas. O verbo concorda com o constituinte que o antecede, o qual é chamado de sujeito da sentença e os constituintes que vêm depois do verbo são os objetos.

Para esse recorte de pesquisa em linguística gerativa, levantemos a seguinte hipótese: O processamento computacional dos sujeitos das sentenças em português opera com deslocamento à direita do verbo para atender a filtro de caso nominativo.

Para explicarmos como esse processamento computacional ocorre nas sentenças do português, traremos uma explicação apresentada por Kenedy (2015, p. 20-23).

Ao descrever as operações formais que geram as estruturas sintáticas na língua portuguesa, por exemplo, devemos entender que as frases que produzimos e compreendemos quando usamos a linguagem são, na verdade, o resultado de um processo computacional. Os gerativistas entendem que a sintaxe de uma língua não funciona naquilo que ouvimos ou escrevemos na superfície de uma dada frase. Eles procuram reconstruir cada operação sintática do processo que lhe deu origem. O conjunto das computações que geram uma frase é chamado de derivação e os sintagmas e as frases que

resultam das operações sintáticas denominam-se representação. Vejamos o exemplo 3 abaixo.

- (3) a. Maria parece gostar de sintaxe.
b. *Parece Maria gostar de sintaxe.

Comparando-se as sentenças (3a) e (3b), observamos que a representação da sentença (3a) é agramatical, pois o sintagma [Maria] é pleno, isto é, possui matriz fonética (não é nulo nem oculto) e está posicionado como sujeito de [gostar], que é verbo na forma infinita. Sabemos que os verbos na forma infinita não conseguem licenciar sujeito. Nesse sentido, ocorre que, nessa frase, existe uma posição sintática livre que licencia uma expressão nominal que atribui a função de sujeito ao verbo flexionado [parece] durante a derivação dessa frase. Isso acontece devido ao sistema computacional que preserva o Filtro de Caso nominativo na operação sintática de deslocamento do constituinte [Maria] para essa posição, antes do verbo [parece]. Podemos ver essa operação na sentença (4).

- (4) [Maria [parece [~~Maria~~] gostar [de [sintaxe]]]]]
-

Observe em (4) que ocorrem o movimento do sujeito da frase [Maria] para uma posição licenciada e o apagamento fonético na posição não licenciada (forma tachada). São operações como essas que tornam as sentenças gramaticais na língua.

Esse exemplo ilustra, de forma simples, a maneira pela qual se postula a existência de operações formais que são necessárias e suficientes para gerar estruturas gramaticais na língua portuguesa. Após o levantamento da hipótese, o gerativista utiliza uma das técnicas mencionadas para aquisição das informações necessárias à pesquisa. Aos poucos, ele observará que uma operação formal advogada pelo pesquisador em sintaxe geral servirá para uma vasta gama de fenômenos linguísticos. Como nós vimos, as regras de movimento são apenas um exemplo de computação no sistema

formal das línguas. Elas se aplicam não só em relação ao licenciamento de sintagmas com Filtro de Caso, mas também em outros fenômenos, como o elemento QU-. Em línguas como o português, elementos QU- são tipicamente deslocados para periferias à esquerda nas frases interrogativas.

2.6 Considerações finais

Os pontos que abordamos sobre a teoria gerativa, neste capítulo, não são suficientes para compreendê-la plenamente, mas vislumbram alguns conceitos necessários à pesquisa em linguística. Se a linguagem for realmente (e parece que é) um órgão da mente/cérebro, então, acreditamos que essa teoria elucida alguns problemas científicos referentes à linguagem e apresenta um caminho de investigação que explica o processamento computacional dos fenômenos linguísticos e a relação da linguagem com outros órgãos da mente/cérebro.

No entanto, consideramos que alguns conceitos operacionais da teoria ficaram de fora ou em stand by para outros capítulos, como as noções de c-comando, s-seleção, c-seleção, projeção sintática, formação de constituintes, espaço derivacional, numeração, SPELL-OUT, etc., as quais são fundamentais para uma compreensão técnica dos processamento computacional das línguas. Outros conceitos aqui apresentados exigiriam maior aprofundamento e exemplificação, como a definição de sistema computacional, sistemas de desempenho, estado inicial da linguagem, que contribuem os conceitos basilares da teoria gerativa.

Todavia, procuramos apresentar o surgimento da teoria gerativa, de modo que o leitor percebesse a importância dela para os estudos da cognição humana e descobrisse a diferença entre o gerativismo e o distribucionalismo. Depois, mostramos os principais conceitos que julgamos necessários à compreensão geral da teoria. Acreditamos que o fulcro da teoria gerativa foi exposto ao pesquisador que pretende iniciar seus estudos no gerativismo, ou seja, diríamos que o tópico sobre princípios teóricos apresentou,

mesmo que de forma geral, as cláusulas *sine qua non* do gerativismo. Com aquelas informações iniciais, o pesquisador foi introduzido às principais noções da teoria gerativa, aos métodos científicos utilizados, às técnicas aplicadas pela teoria para obtenção de resultados e a um pequeno recorte de pesquisa, cujo escopo o investigador poderá ampliar e desenvolver uma pesquisa sobre o movimento de constituintes no português.

REFERÊNCIAS

BARBOSA, José Roberto Alves. **Linguística**: outra introdução. Mossoró: Queima Bucha, 2013. 88 p.

BLOOMFIELD, Leonardo. The use of language. *In*: BLOOMFIELD, Leonardo. **Language**. Londres: George Allen e Unwin, 1935. cap. 2. p. 21-41.

CARVALHO, Cid Ivan da Costa. **Transdutor de estados finitos para conversão de grafema para a pronúncia da variedade linguística potiguar**. 2016. 160 f. Tese (Doutorado em Linguística) – Universidade Federal do Ceará, Fortaleza.

CHOMSKY, Noam. **Novos horizontes no estudo da linguagem e da Mente**. São Paulo: Unesp, 2005. 343 p.

CHOMSKY, Noam. **Linguagem e mente**. São Paulo: Unesp, 2009. 344 p.

CHOMSKY, Noam. **Ciência da linguagem**: conversas com James McGilvray. São Paulo: Unesp, 2014, p. 25-40, 399-403.

CHOMSKY, Noam. **Estrutura sintática**. Petrópolis, RJ: Vozes, 2015.

CHOMSKY, Noam. **Que tipo de criatura somos nós?** Petrópolis, RJ: Vozes, 2018a. 175 p.

CHOMSKY, Noam. **Sobre natureza e linguagem**. 2. ed. São Paulo: Wmf Martins Fontes, 2018b. 243 p.

CHOMSKY, Noam; BERWICK, Robert C. **Por que apenas nós?:** linguagem e evolução. São Paulo: Unesp, 2017.

COSTA, Marcos Antonio. Estruturalismo. *In*: MARTELOTTA, Mário Eduardo. (org.). **Manual de Linguística**. São Paulo: Contexto, 2016, p. 123-126.

JACKENDOFF, Ray. Lexical storage versus online construction. *In*: JACKENDOFF, Ray. **Foundations of language: brain, meaning, grammar, evolution**. Oxford: Oxford University Press, 2002. cap. 6. p. 152-155.

KAYNE, Richard S. Some notes on comparative syntax, with special reference to english and french. *In*: CINQUE, Guglielmo; KAYNE, Richard S. **The Oxford handbook of comparative syntax**. Oxford: Oxford University Press, 2005. p. 3-69.

KENEDY, Eduardo. **Curso básico de linguística gerativa**. São Paulo: Contexto, 2013.

KENEDY, Eduardo. Sintaxe gerativa. *In*: OTHERO, Gabriel de Ávila; KENEDY, Eduardo. **Sintaxe, sintaxes: uma introdução**. São Paulo: Contexto, 2015. Cap. 2. p. 11-26.

LOBATO, Lúcia Maria Pinheiro. **Sintaxe gerativa do português: da teoria padrão à teoria da regência e ligação**. Belo Horizonte: Vigília, 1986. 274 p.

McGILVRAY, James. Introdução. *In*: CHOMSKY, N. **Ciência da linguagem: conversas com James McGilvray**. São Paulo: Unesp, 2014, p. 11-21.

MIOTO, Carlos; SILVA, Maria Cristina Figueiredo; LOPES, Ruth. **Novo manual de morfossintaxe**. São Paulo: Contexto, 2016. 269 p.

PERINI, Mário Alberto. **Sintaxe**. São Paulo: Parábola, 2019.

RAPOSO, Eduardo P. Categorias funcionais na gramática gerativa.

Delta, São Paulo, v. 9, n. 2, p.237-274, jun. 1993.

RIO-TORTO, Graça *et al.* **Gramática derivacional do português**. 2.

ed. Coimbra: Imprensa da Universidade de Coimbra, 2016. 571 p.

Disponível em: <http://dx.doi.org/10.14195/978-989-26-0864-8>. Acesso em: 21 set. 2019.

SILVA, T. C. **Fonética e fonologia do português**. 10. ed. São Paulo:

Contexto, 2014.