

ENSAIO TEÓRICO

HETEROGENEIDADE E HOMOGENEIDADE NAS LÍNGUAS: UMA SÍNTESE HISTÓRICA PARA O SURGIMENTO DA **TEORIA DAS MÚLTIPLAS GRAMÁTICAS** E UMA ANÁLISE SOBRE A ARTICULAÇÃO ENTRE ORAÇÕES EM SENTENÇAS DO PORTUGUÊS BRASILEIRO



OPEN ACCESS

EDITORES

- Miguel Oliveira, Jr. (UFAL)
- René Almeida (UFS)

AVALIADORES

- Elena Preuss (UFG)
- Ricardo de Souza (UFMG)

DATAS

- Recebido: 29/10/2020
- Aceito: 23/11/2020
- Publicado: 14/03/2021

COMO CITAR

KENEDY, Eduardo (2021).
Heterogeneidade e homogeneidade
nas línguas: uma síntese histórica para
o surgimento da teoria das múltiplas
gramáticas e uma análise sobre a
articulação entre orações em
sentenças do Português Brasileiro.
Cadernos de Linguística, v. 2, n. 1, p. 01-
17.

Eduardo KENEDY  

Universidade Federal Fluminense (UFF)

RESUMO

O artigo apresenta o contexto do surgimento e do desenvolvimento da teoria das múltiplas gramáticas (ROEPER, 1999, 2016; AMARAL & ROEPER, 2004, 2014; ABOH, 2015, 2020; KENEDY, 2016, 2018) como um esforço em trazer a realidade heterogênea das línguas naturais para a caracterização gerativista da competência linguística de um falante. Discute o trabalho de Souza (2019) e argumenta que a articulação entre orações em sentenças do português brasileiro representa uma variabilidade relativa ao desempenho linguístico – a saber, a memória de trabalho – e, assim, não encerra um caso de múltiplas gramáticas.

ABSTRACT

This article presents a historical perspective for the rising and development of the multiple grammars theory (ROEPER, 1999, 2016; AMARAL & ROEPER, 2004, 2014; ABOH, 2015, 2020; KENEDY, 2016, 2018), which represents a generativist attempt to characterize the linguistic competence of a native speaker of any language as an intrinsically heterogeneous phenomenon. It also discusses the dissertation of Souza (2019) and argues that the variability of relationships among clauses in Brazilian Portuguese is due to performance factors – namely the working memory – rather than an instance of multiple grammars.

PALAVRAS-CHAVE

Heterogeneidade; Homogeneidade; Múltiplas Gramáticas;
Articulação entre Orações; Português Brasileiro.

KEYWORDS

Heterogeneity; Homogeneity; Multiple Grammars; Clause Relationships;
Brazilian Portuguese.

CONTEXTO DA DISCUSSÃO

Uma das maneiras de contar a história do desenvolvimento das grandes teorias linguísticas dos séculos XX e XXI é descrever como elas vêm tentando conciliar dois aspectos fundamentais – e ao mesmo tempo contraditórios – das línguas naturais: heterogeneidade e homogeneidade. A imensa diversidade existente entre as línguas e dentro de cada uma delas é um fato empírico que um linguista não deve subestimar quando formula teorias e análises mais ou menos homogeneizantes, que procuram caracterizar uma língua ou uma modalidade de língua – ou, ainda, o conhecimento tácito de cada falante sobre essa língua ou modalidade – como um fenômeno coeso e relativamente uniforme. Com efeito, ao longo dos últimos 100 anos, a tensão entre os fatos heterogêneos de uma língua e a ambição homogeneizadora do linguista subjaz a modelos estruturalistas, gerativistas e sociolinguísticos – entre outros (para uma revisão a respeito, cf. BORGES NETO, 2004 e LUCCHESI, 2004). No presente artigo, se apresenta um capítulo específico nessa narrativa: a teoria das múltiplas gramáticas (TMG).

A TMG foi formulada por Roeper no final do século XX e vem sendo desenvolvida por diversos linguistas desde então (cf. ROEPER, 1999, 2016; AMARAL & ROEPER, 2004, 2014; ABOH, 2015, 2020; KENEDY, 2016, 2018) como um modelo formulado para dar conta da representação do conhecimento linguístico (no sentido de Chomsky, 1965) como uma realidade intrinsecamente heterogênea. Em particular, o artigo discutirá um aspecto pontual da diversidade linguística do português brasileiro (PB): a articulação entre orações em sentenças. O propósito dessa discussão é avaliar se as diferentes formas de vinculação oracional em PB caracterizam múltiplas gramáticas na competência linguística dos falantes (cf. SOUZA, 2019) ou se são expressões de uma gramática única, que se distinguem em função de variáveis da ordem do desempenho linguístico.

O artigo está organizado nas quatro seções que se seguem. Na primeira delas, se descreve rapidamente o percurso das principais teorias linguísticas do século XX em suas opções teórico-metodológicas fundamentais que, mais do que tudo, procuravam equilibrar a tensão entre homogeneidade e heterogeneidade no fazer do linguista. Em seguida, se introduz a TMG como uma tentativa de síntese entre o modelo chomskiano de competência linguística e as descobertas da sociolinguística variacionista e dos estudos sobre gêneros textuais a respeito da essência multifacetada das línguas naturais. A terceira seção descreve o fenômeno da articulação entre orações em PB, com ilustrações de casos de parataxe, de hipotaxe e de encaixamento na estrutura de sentenças complexas. Por fim, o artigo põe em discussão a natureza dos distintos modos de vincular orações em PB: são esses um caso de múltiplas gramáticas ou de diferentes formas de desempenho linguístico?

1. BREVE PERCURSO HISTÓRICO

Na obra considerada fundadora da Linguística Moderna – o “Curso de Linguística Geral”, Saussure reconheceu com grande clareza a essência “heteróclita dos fatos da linguagem” (1916/2017, p. 21). Para ele, a cada momento em que um indivíduo produz um enunciado numa língua qualquer, toma corpo um “ato individual de vontade e inteligência” específico (1916/2017, p. 22). Esses atos singulares, no entanto, não poderiam constituir o legítimo objeto a que a então emergente ciência da linguagem deveria se dedicar, pois, na proposta saussuriana, o objeto da linguística (*langue*) deveria ser a realidade abstrata subjacente a cada fala individual e não a realidade concreta dessas mesmas falas. De fato, a *langue*, para Saussure, se apresentava como um fenômeno inteiramente psíquico, tanto nos conceitos mentais suscitados pelos signos, quanto nas imagens acústicas evocadas por cada signo, e, dessa forma, possuiria uma natureza totalmente abstrata e independente dos usos concretos da então nomeada *parole*: a *langue* “é o produto que o indivíduo registra passivamente” (SAUSSURE, 1916/2017, p. 45). Essa dissociação radical entre, de um lado, a *langue* abstrata e psíquica e, de outro, a *parole* concreta e física permitiu que o projeto saussureano caracterizasse as línguas como uma entidade naturalmente homogênea e estática, por oposição ao caráter heterogêneo e dinâmico de cada fala individual, inaugurando uma tradição epistemológica até hoje dominante nos estudos da linguagem.

A linguística estrutural que se seguiu ao projeto saussureano manteve a exclusão da heterogeneidade da fala do centro das preocupações das teorias e das análises linguísticas, mas isso não pôde ser feito sem dificuldade. Artefatos descritivos da fonologia e da morfologia estruturalistas revelavam as perturbações que a realidade da *parole* impunha aos modelos de *langue* nas diferentes línguas: neutralização, debordamento, arquivonema, alofone, alomorfe, variantes livres etc. Cada um desses conceitos confessava haver algum grau de heterogeneidade na *langue*. Além disso, estruturalistas nunca chegaram a ser suficientemente explícitos quanto ao status da sintaxe na arquitetura de um sistema linguístico. O “Curso de Linguística Geral”, por exemplo, não faz menção a estruturas sintáticas. E, como o próprio Chomsky notou, a escola estruturalista como um todo não avançou no conhecimento sobre sintaxe para além do que as gramáticas tradicionais já haviam feito (cf. CHOMSKY, 1986. p. 14). Sendo assim, não se soube em que medida as coerções da *langue* saussuriana atuavam ou não sobre o componente sintático das línguas, o que, dizendo de outra forma, significa que estruturalistas podem ter interpretado a sintaxe como um fenômeno da *parole*, em função de sua natureza heterogênea ou, pelo menos, muito menos homogênea do que a da fonologia.

A virada chomskiana (CHOMSKY, 1957) trouxe a sintaxe para o centro da teorização linguística. Mais do que isso, a então emergente escola gerativista se propunha a descrever

e explicar o conhecimento linguístico individual acerca de uma língua natural qualquer. Se, para Saussure, “um sistema gramatical (...) existe virtualmente em cada cérebro”, mas “não está completo em nenhum” indivíduo, pois “só na massa ele existe de modo completo” (1916/2017, p. 45), para Chomsky uma gramática existiria somente na mente/cérebro de um indivíduo específico, na forma de um conhecimento de língua, a língua-*i* – isto é, um sistema individual, interno e intensional (1986, p. 35). No entanto, esse “indivíduo” chomskiano foi concebido de uma maneira hiperabstrata, num recurso científico que, de certa forma, preservava a teorização homogeneizadora da escola anterior: o “falante-ouvinte ideal, situado numa comunidade completamente homogênea” (CHOMSKY 1965: 83). Dessa forma, a tensão entre homogeneidade e heterogeneidade no fazer da linguística continuaria sendo resolvida com o abafamento da heterogeneidade durante as décadas da hegemonia gerativista.

Na verdade, apesar do grande esforço homogeneizante do paradigma chomskiano, a heterogeneidade do conhecimento linguístico de indivíduos reais sempre se anunciou, de maneira involuntária, na expressão dos “julgamentos de gramaticalidade”, a principal ou única metodologia de análise gerativista. Tais julgamentos frequentemente variam de um falante a outro, de um linguista a outro, ou mesmo de um momento a outro na intuição de um mesmo falante/linguista (COWART, 1997, p. 17). Em geral, gerativistas desprezam a heterogeneidade desses julgamentos e atribuem tal flutuação a fatores dialetais ou lexicais, que são por eles considerados periféricos ou irrelevantes na caracterização do conhecimento de uma língua (cf. SPROUSE, 2007). Todavia, não é raro que a flutuação da impressão da gramaticalidade num único indivíduo ou entre falantes seja levada a sério e, assim, motive o desenvolvimento de modelos gerativistas que confirmam algum grau de heterogeneidade ao conhecimento linguístico. Um exemplo notável nesse sentido é a reinterpretação do parâmetro *pro-drop* para o contexto do PB. Kato (2000) defende a hipótese de que o PB contemporâneo não tenha especificado esse parâmetro como uma questão de tudo ou nada e sustenta que o *pro-drop* nessa língua seja apenas parcial, uma vez que é ainda preservado nos sujeitos não referenciais, mas já desapareceu nos sujeitos referenciais em grande parte das variedades linguísticas brasileiras. Portanto, ao reconhecer a variabilidade presente em uma língua-*i*, Kato (2000) – entre outros sintaticistas, como Holmberg e Roberts (2009) – inaugura uma nova maneira gerativista de equilibrar a tensão entre homogeneidade e heterogeneidade.

Com efeito, a primeira corrente da linguística a deliberadamente enfrentar o desafio de teorizar em meio ao fato da heterogeneidade das línguas foi a sociolinguística variacionista (cf. LABOV, 1972). A proposta laboviana foi inovadora porque interpretava a heterogeneidade linguística não como excepcional ou caótica, mas, antes, como um “sistema”, ou seja, um fenômeno ordenado, regular e previsível. Na sociolinguística, a heterogeneidade passa a ser ressignificada como definidora do próprio sistema linguístico

(LUCCHESI, 2004, p. 163). Em tal sistema, formas e regras categóricas coexistem com formas e regras variáveis controladas por fatores estruturais e socioculturais. Nesse novo paradigma, o desafio do linguista passa a ser descrever a heterogeneidade sistêmica e explicar como ela faculta a aquisição individual e o uso coletivo de uma língua, tal como de fato ocorre em todas as comunidades linguísticas reais.

Entre outros méritos, o enfretamento da heterogeneidade linguística fez com que a escola laboviana pudesse resolver o paradoxo estruturalista acerca da mudança diacrônica. Afinal, se numa dada sincronia a *langue* é homogênea e cada um de seus usos heterogêneos é meramente passivo, conforme argumentou Saussure, quais forças seriam responsáveis pelas transformações diacrônicas registradas em todas as línguas? A contradição saussuriana se desfaz quando Labov propõe que uma língua seja interpretada como uma disputa entre uma grande quantidade de formas e regras variantes (LABOV, 1972, p. 78). Essa competição é a causa da onipresente variação linguística sincrônica e das eventuais mudanças diacrônicas, que acontecem quando uma forma/regra variante provoca o desaparecimento de seus concorrentes – até que novas formas/regras variantes surjam e reiniciem o ciclo da variação à mudança. Além disso, a sociolinguística variacionista é também capaz de explicar a flutuação nos julgamentos de gramaticalidade gerativistas: as formas/regras em competição fazem variar a percepção gramatical de uma frase para a outra, de um indivíduo para o outro, de um momento para o outro. De fato, os testes intuitivos informais gerativistas normalmente não controlam variáveis fonológicas, morfológicas, lexicais ou sociolinguísticas que podem confluir num julgamento de gramaticalidade (cf. COWART, 1997; SPROUSE, 2007), de modo que a atuação de somente uma dessas variáveis já é o bastante para produzir intensa variação individual ou intersubjetiva. Essa possível correlação entre variacionismo e gerativismo foi explorada por linguistas de formação híbrida, numa corrente profícua que se tornou conhecida como sociolinguística paramétrica, da qual o trabalho de Kato (2000) é uma ilustração.

É fato que a sociolinguística trouxe a heterogeneidade para o núcleo da ciência da linguagem. No entanto, a noção de “língua” assumida de maneira tácita ou explícita pelos sociolinguistas apenas parcialmente reconhecia a profusa diversidade das práticas linguísticas humanas. Em verdade, o conceito sociolinguista de “língua” é também estrutural, como o de estruturalistas e gerativista, e equivale ao de “gramática”. Na história da linguística, a noção de língua só seria alterada de maneira significativa a partir das obras seminais de Bakhtin (1953/2003), que só tardiamente viriam a ser conhecidas pelos linguistas do ocidente. Nessas obras, uma língua passa a ser interpretada como um conjunto das práticas discursivas.

A linguística bakhtiniana apontava que os discursos em uma comunidade linguística qualquer são imensamente variáveis e cada um deles ocorre num “gênero” textual específico (cf. BAKHTIN 1953/2003). Um gênero, dizem os bakhtinianos, é uma forma

estável de discurso ancorada num tipo de texto particular, com escolhas lexicais e estruturas gramaticais mais ou menos específicas. Isso quer dizer que um falante-ouvinte real, sociolinguisticamente situado, domina um conjunto de formas e regras específico para cada uma de suas práticas de linguagem: narrar uma história entre amigos, conversar com estranhos, pedir informações a funcionários num estabelecimento, debater em assembleias, escrever mensagens de texto informais, redigir um artigo científico etc. Com a abordagem bakhtiniana, o “uso real” de uma língua passa a pressupor mais do que um sistema linguístico naturalmente variável, mas também uma colossal variedade de gêneros e suas convenções, num universo muito mais heterogêneo do que se havia pensado na tradição estruturalista de Saussure, de Chomsky ou de Labov.

A sociolinguística variacionista e a linguística bakhtiniana nunca se propuseram a explicar a natureza dos estados cognitivos subjacentes a cada ato individual de produção e de compreensão linguística – aquilo que os gerativistas compreendem como uma língua-*i*. Da mesma forma, o gerativismo chomskiano mais ortodoxo jamais se importou com o fato da heterogeneidade dos usos linguísticos. Dessa maneira, os estudos de cada uma dessas correntes se desenvolveram de maneira independente durante muitos anos. Porém, para o progresso da ciência da linguagem, integrar de alguma forma o conhecimento produzido por essas diferentes áreas da linguística é, ao mesmo tempo, uma necessidade epistemológica e um desafio teórico-metodológico. Isso só começou a ser esboçado nos trabalhos de Roeper ao fim do século XX.

2. MÚLTIPLAS GRAMÁTICAS

Na história narrada neste artigo, Roeper (1999) se destaca como o linguista que formulou uma teoria ambiciosa, a qual trazia ao centro do empreendimento gerativista aspectos da heterogeneidade natural de todas as línguas humanas. Não obstante, convém explicitar que, ao apresentar e desenvolver a TMG, Roeper nunca mencionou qualquer preocupação com a sociolinguística ou com os gêneros discursivos. Sua proposta ao longo dos anos (ROEPER, 1999, 2016; AMARAL & ROEPER, 2004, 2014) se manteve nos limites do arcabouço teórico gerativista. Foi apenas com a reinterpretação de Kenedy (2016, 2018) que a TMG passou a ser deliberadamente articulada às noções de variável sociolinguística e de gênero textual.

Em sua proposta original, Roeper recorreu a evidências sobre a aquisição da linguagem para defender a necessidade de representar uma língua-*i* na forma de múltiplas gramáticas. O autor destacou que, em qualquer língua, o input apresentado a bebês e crianças é extremamente variável, no sentido de que informações gramaticais incompatíveis com uma única gramática caracterizam os atos de fala reais de uma

comunidade linguística qualquer. Por exemplo, uma criança que adquire uma modalidade do inglês norte-americano eventualmente será exposta a excertos de fala que indicam se tratar de uma língua [- pro-drop] – quando sujeitos são sempre preenchidos, mesmo se expletivos: *I know him, There is a cat here, It is hot in here*. Simultaneamente, essa mesma criança analisará dados de fala que veiculam informação contrária, indicativa de que a língua em aquisição possui o parâmetro [+ pro-drop] – e então sujeitos de todas as categorias podem ser nulos: *seems nice, raining now?, wanna go out*. De acordo com Roeper, contradições dessa natureza se espriam por todos os domínios da gramática de uma língua: ordem, movimento, caso, concordância, encaixamento, adjunção etc. Ora, de um ponto de vista gerativista ortodoxo, tais contradições deveriam impedir a aquisição da língua, afinal como um parâmetro poderia ser formatado pelas crianças, se a informação do ambiente é compatível com diferentes configurações na gramática? Como uma regra gramatical poderia ser adquirida, se a informação do ambiente indica que ela ora se aplica e ora não?

A solução apresentada por Roeper (1999) sustentava que a criança deveria adquirir múltiplas gramáticas em paralelo. Para o autor, cada uma dessas gramáticas contém os parâmetros e as regras específicos de uma modalidade da língua (reinterpretada por Kenedy, 2016, 2018 como variedade sociolinguística e/ou gênero discursivo). Por exemplo, o parâmetro pro-drop em inglês seria formatado como positivo em certos registros da língua, mas como negativo para muitos outros registros. Essas configurações múltiplas, espalhadas por todos os módulos do sistema linguístico, dariam origem a um conjunto de “minigramáticas”, cada qual usada numa prática discursiva específica. Para Roeper, tal mosaico de pequenas gramáticas caracterizaria a competência linguística de um falante nativo e, dessa forma, representaria a solução gerativista capaz de trazer ao centro da teorização sobre uma língua-i o fato da heterogeneidade linguística.

Kenedy (2016, 2018) sustentou que a intensa diversidade linguística existente no Brasil oferece uma oportunidade singular para a aplicação da TMG. Segundo o autor, os diversos gêneros da escrita e da fala mais formal e culta do PB demandam computações gramaticais muito diferentes daquelas adquiridas espontaneamente nas variedades vernaculares dessa língua, também em seus diversos gêneros de interação informal oral face a face. Kenedy (2016: 196-167) argumenta que os vernáculos do PB estão num extremo de um contínuo de heterogeneidade linguística em que, entre outros fatores, predominam estruturas de tópico > comentário (cf. PONTES, 1987; NEGRÃO, 1990), sujeitos são preenchidos (cf. DUARTE, 1995; KATO, 2002), há fraca concordância verbo-nominal (cf. SCHERRE, 1993, 1994) e pronomes acusativos de 3ª pessoa são tônicos ou nulos (cf. CYRINO, 1997), por contraste ao outro extremo do continuum, a escrita de gêneros mais formais, no qual, além de outras variáveis, predominam estruturas sujeito > predicado, sujeitos são nulos, há forte concordância verbo-nominal e pronomes acusativos de 3ª

pessoa são clíticos. Nesse cenário, brasileiros em processo inicial de literacia (considerando “literacia” como proficiência na produção e na compreensão de textos nos gêneros mais elaborados e prestigiados da cultura escrita) possuíam um número restrito de gramáticas, quando comparados aos brasileiros plenamente letrados, que, ao longo dos anos de literacia, teriam adquirido, para além de sua gramática vernacular, uma série de outras gramáticas, responsáveis pela performance nos diversos gêneros de fala e de escrita sociolinguisticamente mais prestigiados.

3. ARTICULAÇÃO ENTRE ORAÇÕES

Em suas investigações experimentais sobre as habilidades de leitura e escrita de calouros universitários brasileiros, desenvolvidas a partir das hipóteses de Kenedy (2016, 2018), De Souza e Kenedy (2019) identificaram que a articulação entre orações na composição de sentenças era um dos principais fatores que dificultavam o desempenho proficiente da leitura desses estudantes, assim como era essa uma das características que mais flagrantemente diferenciavam a escrita desses alunos e os textos dissertativos-argumentativos alvos do ensino escolar – utilizados em exames como Pisa e ENEM. Esses achados foram aprofundados na dissertação de Souza (2019), desenvolvida sob a orientação de Kenedy. A autora pesquisou as habilidades de redação de estudantes de ensino médio de escolas públicas do RJ em situação de vulnerabilidade social, confirmou as análises de De Souza e Kenedy (2019) e levantou a hipótese de que os diferentes mecanismos de articulação entre orações em sentenças pudessem configurar um caso de múltiplas gramáticas em PB. Essa interpretação foi motivada pelos estudos de Givón (1990) acerca da gramaticalização da vinculação oracional.

Givón, de uma perspectiva translinguística e diacrônica, asseverou que a articulação entre orações se dá de três maneiras principais: parataxe, hipotaxe e encaixamento. Segundo o autor (1990/2003, p. 133), a parataxe se caracteriza como um mecanismo de simples justaposição entre orações, no qual duas ou mais orações sintaticamente independentes se seguem na linearidade da frase, sem marcas linguísticas que explicitem a relação gramatical ou proposicional entre elas, de modo que tais relações devem ser inferidas pragmaticamente. Souza (2019, p. 56) ilustra um caso de parataxe na redação de seus estudantes, reproduzido em (1). É importante notar que a segunda oração se segue linearmente à primeira, com a qual mantém uma relação de continuidade de tópico discursivo sem vínculos sintáticos explícitos.

(1) (...) “para tudo a uma solução, nem tudo está no poder dos homens”.

Já a respeito da hipotaxe, Givón indica que articulações desse tipo se dão por meio de um vínculo caracteristicamente adverbial entre orações: uma oração é um adjunto sintático de outra. Na articulação hipotática, argumenta (GIVÓN, 1990/2003, p. 134), normalmente ocorre a mediação de um conectivo que explicita a relação lógica entre o conteúdo das proposições veiculadas por cada oração (concessão, causa, consequência, tempo etc.). Souza (2019, p. 55-56) apresenta o seguinte caso de hipotaxe entre as redações que analisou. É preciso notar que a segunda oração é um adjunto adverbial da oração anterior, cujo conector “para” explicita a relação de finalidade estabelecida entre o conteúdo da segunda oração em relação ao da primeira.

(2) (...) “As mulheres exigem um pouco mais de necessidade em seus próprios objetivos, para elas terem um pouco mais de independência”.

Por fim, Givón sustenta que o encaixamento encerra a articulação entre uma oração que é um argumento de um constituinte da outra (por exemplo, um complemento, um sujeito ou um predicativo). Nesse tipo de articulação, um conectivo, se houver, será um elemento complementador esvaziado de conteúdo interpretativo (GIVÓN, 1990/2003, p. 134-135). Isso é exemplificado por Souza (2019, p. 56) na frase que se segue. É notável que a segunda oração é um complemento do verbo “dizer”, presente na primeira oração, e que o elemento “qu-” é um conector desprovido de valor semântico ou pragmático.

(3) “Podemos dizer que vivemos num mundo machista e opressor”.

O mais importante da análise de Givón é a sua interpretação de que cada uma dessas formas de articulação entre orações representaria um momento específico no processo de gramaticalização de uma língua (GIVÓN, 1990/2003, p. 130). Para o autor, a parataxe representaria uma gramática ainda embrionária, restrita à justaposição e com grande dependência do contexto e da informação pragmática para determinar tanto a ordem de ocorrência de cada oração, como a interpretação do valor comunicativo a ser estabelecido entre elas. Givón propõe que essa gramaticalização mais elementar seria típica das primeiras línguas no processo de evolução da linguagem (filogênese), bem como, até o presente, seria manifestada logo nos primeiros anos de desenvolvimento da linguagem na vida de uma criança (ontogênese). A hipotaxe, continua (GIVÓN, 1990/2003, p. 138), definiria uma gramática mais estruturada, que restringiria a posição linear de uma oração relativamente a outras e determinaria a sua interpretação com o uso de um conectivo visível, abandonando, dessa forma, a dependência total de um contexto pragmático de interlocução. Essa gramática hipotática seria, para Givón, mais tardia na

evolução das línguas e representaria estágios mais desenvolvidos nos processos de aquisição da linguagem. Por último, Givón caracteriza o encaixamento como o estágio mais avançado da gramaticalização das relações entre orações, no qual uma oração passa a ser parte integrante de outra. Esse seria o último estágio da gramaticalização da vinculação entre orações na história das línguas e o mais tardio no desenvolvimento linguístico de um indivíduo particular – e isso não se daria por acaso, pois, de acordo com Givón, “a ontogênese recapitula a filogênese” durante a aquisição de uma língua pela criança (GIVÓN, 1995, p. 93).

Ao levar a sério as hipóteses de Givón (1990), Souza (2019) revelou que a gramática subjacente às redações dos 17 alunos concluintes de ensino médio por ela estudados deveria ser interpretada como eminentemente paratática. De fato, nos textos dissertativos-argumentativos produzidos por esses estudantes, das 225 articulações oracionais catalogadas, 43% ocorriam por meio da parataxe, 30% por meio da hipotaxe e 27% por meio de encaixamento. Esse quadro é muito diferente das articulações identificadas por Souza na redação por ela tomada como controle, redigida por um aluno de pós-graduação, que manifestou 15 articulações de orações, das quais 43% ocorriam via encaixamento, 37% via hipotaxe e 20% via parataxe. Diante desses dados, Souza defendeu a hipótese de que a gramática de indivíduos plenamente letrados em PB (sua redação controle) disponha de recursos hipotáticos e de encaixamento que são apenas elementares ou francamente inexistentes na gramática dos indivíduos ainda em processo de literacia (suas 17 redações).

É provável que o contexto encontrado por Souza (2019) se reproduza para além da produção dos estudantes analisados em sua dissertação, tal como se pode inferir a partir das milhões de redações do Exame Nacional do Ensino Médio. No exame do ano de 2017 (cf. INEP, 2017), por exemplo, do qual participaram mais de 6 milhões de estudantes, 70% das redações apresentaram notas inferiores à pontuação de uma redação considerada apenas mediana (isto é, a que atingisse 500 pontos em mil). Uma amostra das redações medianas publicizadas pela imprensa brasileira indica os seguintes dados médios acerca da vinculação entre orações: 8 articulações oracionais por redação, com predominância da parataxe (45%), seguida da hipotaxe (35%) e do encaixamento (20%). Em contraste, as redações com as notas mais altas no ENEM daquele ano (1% do total, com pelo menos 900 pontos em mil) apresentavam em média 17 articulações, com 58% de hipotaxe, 37% de encaixamento e apenas 5% de parataxe.

Os dados de Souza (2019) e do ENEM indicam que a maioria dos estudantes brasileiros de fato possui grande dificuldade em empregar os recursos de articulação oracional típicos de textos dissertativos-argumentativos. De acordo com Souza (2019), isso aconteceria porque os diferentes gêneros textuais praticados no PB demandam gramáticas específicas para a articulação entre orações. Segundo a autora, os gêneros

textuais orais e espontâneos, de interação face a face, requereriam gramáticas paratáticas, enquanto os gêneros escritos mais formais e planejados exigiriam gramáticas eminentemente hipotáticas e encaixadas. Os estudantes analisados por Souza e os excertos medianos do ENEM estariam, assim, revelando a permanência de uma gramática paratática em gêneros do discurso nos quais ela é improdutiva. Tal limitação ocorreria porque os estudantes teriam falhado em adquirir gramáticas mais hipotáticas e encaixadas durante os anos de sua escolarização.

4. GRAMÁTICA OU MEMÓRIA DE TRABALHO

O modelo de Givón (1990) é interessante porque afirma que um estágio mais avançado na gramaticalização da articulação entre orações necessariamente contém os estágios anteriores. Dessa maneira, línguas que possuem encaixamento entre orações possuirão também hipotaxe e parataxe, da mesma forma que línguas que possuem hipotaxe, mas não encaixamento, possuirão parataxe. Ao se considerar essa hipótese givoniana, a heterogeneidade do PB descrita por Souza (2019) pode ser interpretada não como uma questão de múltiplas gramáticas, mas, antes, como um problema da ordem do desempenho linguístico, relativo à maior ou à menor fluência individual em um gênero textual específico.

Essa interpretação é importante porque, nos dados de Souza e do ENEM citados, a ocorrência de cada tipo de articulação entre orações é medida em termos de grau – e não nos termos absolutos de sua ausência ou presença. Com efeito, todas as redações analisadas por Souza e divulgadas pelo INEP manifestaram os três tipos de articulação conforme descritos por Givón (1990), logo todos os falantes do PB examinados possuem, em alguma medida, gramáticas para a parataxe, para a hipotaxe e para o encaixamento. Sendo assim, se não dizem respeito a múltiplas gramáticas do PB, o que os dados de Souza (2019) podem significar no contexto da tensão entre “homogeneidade X diversidade” discutida no presente artigo?

Em primeiro lugar, é preciso notar que a hipotaxe se distingue da parataxe e do encaixamento por ser capaz de produzir rupturas na linearidade das informações apresentadas numa sentença. Por exemplo, na frase ilustrada em (2), abaixo reproduzida como (4a), a oração adverbial pode ser posicionada antes de sua oração matriz de referência, como ocorre em (4b), ou mesmo de maneira interpolada nessa oração (4c).

(4a)(...) “As mulheres exigem um pouco mais de necessidade em seus próprios objetivos, para elas terem um pouco mais de independência”.

(4b)(...) “Para elas terem um pouco mais de independência, as mulheres exigem um pouco mais de necessidade em seus próprios objetivos”.

(4c)(...) “As mulheres exigem, para elas terem um pouco mais de independência, um pouco mais de necessidade em seus próprios objetivos”.

Essa distribuição não linear dos conteúdos proposicionais das orações numa dada sentença impõe à memória de trabalho do falante-ouvinte uma demanda específica e mais custosa em relação a uma disposição linear como a da parataxe, sobretudo quando se considera uma sequência de muitas frases na fluência natural de um texto (cf. MAIA, 2019, p. 78). Durante o uso linguístico, construir, frase a frase, um modelo mental a partir de informações que se sucedem linearmente é, para a memória, muito menos custoso do que conectar informações espalhadas em diferentes posições lineares, conforme atestam diversas pesquisas experimentais em psicolinguística (cf. GIBSON, 2000). Sendo assim, uma articulação como a em (4a) é cognitivamente menos custosa do que a de (4b) ou (4c), logo uma sequência de frases como (4a) imporá menos demanda sobre a memória de trabalho do que uma sequência de frases como (4b) ou (4c).

Em segundo lugar, não se deve ignorar que a interação oral face a face, nos gêneros de discurso mais informais e familiares, ocorre tipicamente com proposições veiculadas de maneira linear, algo que favorece à ocorrência da parataxe (cf. SCHOLLES & WILLIS, 1987). Por sua vez, na escrita formal e nos gêneros orais mais próximos dessa escrita, o conteúdo proposicional de uma oração pode ser anteposto ao de outra, ou mesmo a ela interpolado, a depender do tipo particular de efeito que essa linearização pode provocar no discurso. Essas anteposições e interpolações são especialmente produtivas nos textos dissertativos-argumentativos, que apresentam graus de planejamento e monitoramento muito superiores ao da fala espontânea dos gêneros mais informais e familiares e, por isso, sobretudo na modalidade escrita, favorecem usos linguísticos altamente demandantes da memória de trabalho.

É muito importante destacar que, entre os casos de hipotaxe catalogados por Souza, bem como entre os exemplos de hipotaxe das redações medianas do ENEM do ano de 2017, não ocorre sequer uma única relação sintática não local entre orações, ou seja, em todos os casos de hipotaxe dos textos analisados, uma oração adverbial sempre sucede a sua oração matriz de referência, de modo que as interpretações locais e lineares típicas da parataxe se preservam também nos usos hipotáticos dos participantes da pesquisa de Souza (2019) e do ENEM. De fato, em todas as redações citadas ocorrem estruturas como (4a), entretanto em nenhuma delas se registrou algo como (4b) ou (4c). O mesmo deve ser dito a respeito das estruturas encaixadas: em nenhuma das redações, elas provocavam computações linguísticas não locais, isto é, nos dados de Souza e do ENEM, uma oração

encaixada sempre mantinha uma relação linearmente local com o constituinte de referência na oração matriz. Por seu turno, a redação tomada por Souza como controle e as redações ilustrativas das notas mais altas no ENEM manifestavam em média 35% de articulações sintáticas não locais, com orações hipotáticas antecipadas, interpoladas ou separadas de suas orações de referência por diversos tipos de material linguístico interveniente. Portanto, entre essas redações, pouco mais de um terço das articulações oracionais se davam como em (4b) ou (4c), muito embora a maioria das vinculações entre orações acontecesse de modo local como em (4a).

Tais dados parecem evidenciar que a variável determinante na heterogeneidade da articulação oracional em PB não é a gramática internalizada pelo falante, mas sim a sua capacidade de construir e interpretar articulações sintáticas de maneira não local, um fenômeno do desempenho linguístico controlado pela memória de trabalho. Articulações sintáticas à distância demandam maior poder computacional da memória de trabalho e, por isso mesmo, essa habilidade psicolinguística normalmente não se desenvolve de maneira espontânea e imediata a partir das práticas discursivas informais de interações orais face a face (cf. SCHOLES & WILLIS, 1987). Na verdade, assim como acontece com as demais especificidades dos gêneros discursivos mais planejados e formais, o domínio de articulações oracionais hipotáticas não locais requer muito treinamento em língua escrita, algo que tem lugar no curso dos anos de literacia (cf. SCHOLES & WILLIS, 1987, p. 81).

Diante disso, é possível sustentar que os estudantes brasileiros citados por Souza (2019) e representados pelos excertos do ENEM possuam, em suas gramáticas, articulações oracionais de todos os tipos, mas não tenham a sua memória de trabalho treinada no nível necessário à compreensão ou à produção da articulação hipotática ou encaixada não linear demandada pelos textos dissertativos-argumentativos prototípicos. Nesse cenário, o que os dados sobre articulação entre orações em PB revelariam seria uma distribuição desigual, possivelmente diglósica, de diferentes práticas discursivas: num extremo, estariam os gêneros de discursos socialmente prestigiados e explorados em exames como ENEM e Pisa e, no outro extremo, estariam os gêneros da fala e da escrita menos prestigiados e mais próximos da oralidade vernacular não letrada. Como as práticas dos gêneros mais prestigiados impõem à memória de trabalho recursos que não são explorados nas relações mais lineares típicas das práticas vernaculares, indivíduos não plenamente letrados possuiriam severas limitações em suas habilidades de produzir ou compreender textos naquelas modalidades.

5. PARA CONCLUIR

Um século depois de Saussure, a linguística contemporânea tem se mostrado bastante madura ao teorizar em meio à heterogeneidade exuberante das línguas naturais. É fato que o falante-ouvinte ideal chomskiano ainda sobrevive no fazer homogeneizante dos linguistas, mas abordagens como a da TMG permitem o desenvolvimento de teorias sobre a competência linguística muito mais fiéis ao mundo real, em sua grande diversidade de modalidades sociolinguísticas e de gêneros discursivos. O presente artigo reconhece e louva esforços como o da TMG e, ao mesmo tempo, chama a atenção para o fato de que certos aspectos heterogêneos nas habilidades linguísticas de um falante-ouvinte real podem decorrer de variáveis psicolinguísticas não gramaticais, dentre as quais a memória de trabalho mais ou menos bem treinada para produzir e compreender estruturas de dependência sintática não local é um bom exemplo.

O estudo de Souza (2019) tal como aqui reinterpretado destaca que a diversidade linguística do PB é também um retrato da desigualdade de oportunidades de que o Brasil ainda não conseguiu se libertar, pois as práticas discursivas mais prestigiadas e socialmente importantes, seja na modalidade oral ou na escrita, não integram consistentemente o inventário das habilidades linguísticas da maioria dos estudantes brasileiros, ao menos a partir do que se pode deduzir pela gigante quantidade de redações inframedianas do ENEM. Como se argumenta neste artigo, os processos de literacia promovem nos estudantes não somente o aprendizado dos traços do léxico e da gramática necessários à proficiência nos gêneros textuais mais formais e influentes na língua, mas também permitem o desenvolvimento das habilidades cognitivas gerais essenciais ao processamento do texto escrito, como o treinamento da memória de trabalho. Para poucos brasileiros, o acesso a esses ambientes de literacia é pleno, todavia para a maioria da população ele é pobre ou nulo (cf. KENEDY, 2018, 2016). O resultado dessa desigualdade parece ser uma situação diglósica em termos de práticas discursivas.

A escola pública brasileira ainda é o espaço em que a população de origem trabalhadora tem acesso a processos de literacia. Nesse sentido, é importante que a natureza e a dinâmica da heterogeneidade linguística do PB sejam plenamente conhecidas de professores e demais profissionais da educação. Promover em sala de aula o desenvolvimento das variáveis linguísticas e cognitivas mais gerais relevantes para a literacia dos estudantes é indispensável ao ainda distante sucesso da formação escolar básica neste país.

6. AGRADECIMENTOS

À Direção da ABRALIN, pela iniciativa da “ABRALIN ao Vivo – *Linguists On-Line*”, que deu origem às discussões apresentadas neste artigo.

Aos críticos da apresentação “A teoria das múltiplas gramáticas: o caso da articulação entre orações no PB”, veiculada no canal da ABRALIN no Youtube e, 6/6/2020, em especial a Marcus Maia, Eloisa Pilati, Leonor Scliar-Cabral, Márcio Leitão, Luciana Sanchez Mendes, Sabrina Lopes, Luiz Amaral e todos os estudantes do Laboratório de Psicolinguística Experimental (UFRJ), do Laboratório de Processamento Linguístico (UFPB) e do Grupos de Estudos e Pesquisa em Linguística Teórica e Experimental (UFF): o presente texto procurou contemplar as objeções e as sugestões por vocês apresentadas em decorrência daquela *live* da ABRALIN.

Aos pareceristas deste artigo, cujas valiosas sugestões foram todas incorporadas à versão final do texto.

Ao CNPq, pela Bolsa de Produtividade em Pesquisa que permitiu os anos de pesquisa sobre múltiplas gramáticas e sobre processamento psicolinguístico de sentenças.

REFERÊNCIAS

- ABOH, E. Lessons from neuro-(a)-typical brains: universal multilingualism, Code-Mixing, Recombination, and Executive Functions. *Frontiers in Psychology*, 11, 2020.
- ABOH, E. *The emergence of hybrid grammar: contact and change*. Cambridge: CUP, 2015.
- AMARAL, L. & ROEPER, T. Multiple Grammars and second language representation. *Second Language Research* 30(1), 2014.
- AMARAL, L. & ROEPER, T. Why minimal multiple rules provide a unique window into UG and L2. *Second Language Research*, 97, 2004.
- BAKHTIN, M. Os Gêneros do discurso. In: *Estética da criação verbal*. Martins Fontes, SP, 2003 – edição original de 1953.
- BORGRES NETO, J. *Ensaio de filosofia da linguística*. Parábola: SP, 2004.
- CHOMSKY, N. *Knowledge of language: its nature, origin and use*. NY: Praeger, 1986.
- CHOMSKY, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- CHOMSKY, N. *Syntactic structures*. NY: The Hague Mouton, 1957.
- COWARD, W. 1997. *Experimental syntax: applying objective methods to sentence judgments*. London: Sage Publications.
- CYRINO, S. O objeto nulo no português do Brasil e no português de Portugal. *Boletim da ABRALIN*, n. 25, 2001, pp. 173-81.
- DE SOUZA, J. A. & KENEDY, E. Repensando as habilidades de leitura no ensino superior sob a ótica da psicolinguística experimental. *Revista da ABRALIN*, v. 16, n. 1, 2019, pp. 350-375.

- DUARTE, M. E. L. *Perda do princípio "Evite Pronome" no português brasileiro*. Tese de Doutorado (Doutorado em Linguística). Campinas: UNICAMP, 1995.
- GIBSON, E. The dependency locality theory: a distance-based theory of linguistic complexity. In MARANTZ, A.; MIYASHITA, Y. & O'NEIL, W. (Ed.). *Image, language, brain*. papers from the first mind articulation project symposium, Cambridge, MA: MIT Press, p. 95-126, 2000.
- GIVÓN, T. *Functionalism and grammar*. Amsterdam: John Benjamins, 1995.
- GIVÓN, T. *Syntax: a functional-typological introduction*. Amsterdam: John Benjamins, 1990.
- HOLMBERG, A.; ROBERTS, I.; Introduction: parameters in minimalist theory. In: BIBERAUER, T. et al. *Parametric variation: null subjects in minimalist theory*. Cambridge University Press, 2009.
- INEP. *Resultados dos participantes de 2017*. Disponível em: inep.gov.br/educacao_basica/enem/downloads/2017/apresentacao_resultados_enem2017.pdf. Acessado em 28/10/2020.
- KATO, M. A. A gramática do letrado: questões para a teoria gramatical. In: MARQUES, M. A. et al. (Orgs.). *Ciências da linguagem: trinta anos de investigação e ensino*. Braga: CEHUM, 2002.
- KATO, M. The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. *Romania Nova*. Campinas, 2000.
- KENEDY, E. O problema do analfabetismo funcional no Brasil sob uma análise psicolinguística. In: MAIA, M. (Org.). *Psicolinguística e educação*. Campinas: Mercado de Letras, 2018, p. 81-102.
- KENEDY, E. O status da norma culta na língua-i dos brasileiros e seu respectivo tratamento na escola: algumas contribuições de estudos formalistas à educação. In: GUESSER, S. *Linguística: pesquisa e ensino*. Boa Vista: EDUFRR, 2016. pp. 173-196.
- LABOV, W. *Sociolinguistic patterns*. Philadelphia, University of Pennsylvania Press, 1972.
- LUCCHESI, D. *Sistema, mudança e linguagem: um percurso na história da linguística moderna*. Parábola: SP, 2004.
- MAIA, M. Rastreamento ocular de períodos compostos e consciência sintática. In: MAIA, M. (Org.). *Psicolinguística e metacognição na escola*. SP: Mercado de Letras, 2019.
- NEGRÃO, E. V. *O português brasileiro: uma língua voltada para o discurso*. SP: USP, 1990.
- PONTES, E. *O tópico no português do Brasil*. Campinas: Pontes, 1987.
- ROEPER, T. Multiple Grammars and the logic of learnability in second language Acquisition. *Frontiers in Psychology*, 7-14, 2016.
- ROEPER, T. Universal bilingualism. *Bilingualism, Language and Cognition* (2), 169-186, 1999.
- SAUSSURE, F. *Curso de linguística geral*. SP: Cultrix, 2017 – original de 1916.
- SCHERRE, M. M. P. & NARO, A. J. Duas dimensões do paralelismo formal na concordância de número no português popular do Brasil. *D.E.L.T.A.*, v. 9, n.1, p. 1-14, 1993.
- SCHERRE, M. M. P. Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua Portuguesa* (RILP), v. 12, p. 37-49, 1994.
- SCHOLES, R. & WILLIS, B. Age and education in oral language skills. *Developmental Neuropsychology*, 3, 239 -253, 1987.
- SOUZA, M. *Múltiplas gramáticas: um estudo sobre a diglossia entre a variedade da escrita culta e os vernáculos do português brasileiro*. Dissertação (Mestrado em Estudos da Linguagem), Instituto de Letras, Universidade Federal Fluminense. Niterói, 2019.
- SPROUSE, J. *A Program for Experimental Syntax: Finding the relationship between acceptability and grammatical knowledge*. Doctoral Dissertation (PhD in Linguistics). University of Maryland. 2007.