

# Notas sobre o *corpus* em estudos da linguagem: a perspectiva fractal e a Técnica da Saturação Teórica

*Notes on the corpus in language studies: the fractal perspective and the Theoretical Saturation Technique*

A. Ariadne DOMINGUES ALMEIDA

Universidade Federal da Bahia  
ada.domingues@gmail.com.br



**Resumo:** Neste artigo, a linguagem e o *corpus*, constituído para estudá-la, são compreendidos como fractal — objeto que mantém sua forma invariável, quando medido em qualquer escala; que é obtido por processos recursivos e que apresenta auto-similaridade, complexidade infinita e dimensão. A Técnica da Saturação é uma ferramenta conceitual que, amparada por essa perspectiva fractal da linguagem, sustenta a constituição do *corpus*. Em face destas concepções e, com base nas discussões viabilizadas pela Teoria da complexidade (Capra, 2006 [1996]; Morin, 2009 [1999]; Maturana; Varela, 2001 [1984]), aqui, é proposto, no plano metodológico, o emprego da Técnica da Saturação (Glaser; Strauss, 1967; Fontanella, Ricas; Turato, 2007), visando à constituição de *corpus* no âmbito dos estudos da linguagem e, em particular, de pesquisas empreendidas em Linguística Cognitiva, demonstrando a sua aplicabilidade através do fenômeno da conceptualização, a partir dos trabalhos de Santana (2019) e Duque (2020). Os resultados alcançados por esses autores demonstram a validade dessa proposta teórico-metodológica para o desenvolvimento de estudos acerca dos fenômenos da linguagem.

**Palavras-chave:** linguagem; fractal; *corpus*; Técnica da Saturação.

**Abstract:** In this article, language and the *corpus* constituted to study it are understood as a fractal — an object that maintains its invariable form, when measured on any scale; something obtained by recursive processes and that presents self-similarity, infinite complexity and dimension. The Saturation Technique is a conceptual tool that, supported by this fractal perspective of language, supports the constitution of the *corpus*. In view of these conceptions and, based on the discussions made possible by the Theory of Complexity (Capra, 2006 [1996]; Morin, 2009 [1999]; Maturana; Varela, 2001

[1984]), here, at the methodological level, the use of the Saturation Technique is proposed (Glaser; Strauss, 1967; Fontanella, Ricas; Turato, 2007), aiming at the constitution of a *corpus* in the field of language studies and, in particular, of research undertaken in Cognitive Linguistics, demonstrating its applicability through the phenomenon of conceptualization, from the works of Santana (2019) and Duque (2020). The results achieved by these authors demonstrate the validity of this theoretical-methodological proposal for the development of studies on the phenomena of language<sup>1</sup>.

**Keywords:** language; fractal; *corpus*; Saturation Technique.

---

<sup>1</sup> Este texto oferece reflexões atreladas ao meu Pós-Doutorado desenvolvido na Universidade Federal do Rio Grande do Norte.

## 1 INTRODUÇÃO

No âmbito dos estudos linguísticos, pelo menos, dois percursos têm sido traçados para estabelecer diálogos com a Teoria da Complexidade (TC). Por um lado, é possível fazer referência a algumas abordagens do fenômeno linguageiro pelo viés desse paradigma científico, a exemplo da Linguística do Caos ou Linguística Catastrófica, campo do conhecimento que, paulatinamente, tem publicado resultados das investigações já empreendidas, a exemplo do artigo de Zamoro Aguilar (2012) que recorre à Teoria do Caos para discutir a comunicação verbal. Por outro lado, também, são feitas algumas aproximações entre a TC e outras vertentes da Linguística, a exemplo dos diálogos produzidos com a Linguística Aplicada, a fim de promover, entre outras, reflexões sobre a aquisição da linguagem, como as feitas por Paiva (2005; 2011; 2016).

Aqui, seguirei esse segundo percurso, mas com enfoque em interações entre a TC e a Linguística Cognitiva (LC). Apesar de não serem abundantes, já não são tão novas as pontes que podem ser construídas entre esses dois campos do conhecimento. Assim sendo, em 1994, Bernárdez promove uma discussão acerca das aproximações possíveis entre essas áreas do saber. Além dele e, também, citado por ele, Petitol (1991) empregou a TC, com a finalidade de formalizar modelos da LC, e, ainda, Brandt (1992) propôs, com base nos estudos de Talmy e Sweetser, uma aproximação entre essas duas perspectivas da investigação científica e, com isto, estabeleceu um alicerce sólido para estudar a linguagem. Além desses autores, outros a exemplo de Teixeira (2004), Fernández Jaén (2008), Paiva (2010), Almeida Júnior (2013) propuseram migrações conceituais da TC para a LC.

No texto *A tessitura do conhecimento: o corpus na construção de estudos semânticos sócio-histórico-cognitivos* (Almeida, 2020a), pressupostos da Matemática dos fractais possibilitaram-me discutir como a compreensão do *corpus* como fractal pode contribuir para o desenvolvimento de pesquisas semânticas de natureza qualitativa. Aqui, busco ampliar o debate, de sorte que proponho refletir sobre a concepção de linguagem como fractal, retomo a noção do *corpus* também como um fractal, para enfim abordar efetivamente a Técnica da Saturação Teórica, que, no artigo anterior, foi apenas indicada como um possível caminho para constituição do *corpus* no âmbito das investigações científicas em Semântica Cognitiva (SC). Agora, reflito sobre os contributos dessa ferramenta conceitual para a realização dos trabalhos sobre o fenômeno da conceptualização. Para iniciar o debate, realizo, na próxima seção, um breve

apanhado da concepção de fractal para, em seguida, focar essa noção no plano da linguagem.

## 2 A TEORIA DOS FRACTAIS: UMA MATEMÁTICA PARA FALAR DA LINGUAGEM

Antes de abordar a linguagem pela perspectiva da Teoria dos Fractais, farei algumas considerações sobre a própria noção de fractal e da sua matemática. Inicialmente, destaco, a partir de Capra (2006 [1996]), que a Teoria dos Fractais, bem como a Teoria do Caos, vertentes da Teoria dos Sistemas Dinâmicos, enfim, da Teoria da Complexidade, é uma teoria matemática com conceitos e técnicas aplicados a uma série de fenômenos, mas não é uma teoria dos fenômenos físicos.

Além disto, observo que a Teoria dos Fractais — Geometria Fractal — foi gestada independente da Teoria do Caos, entre as décadas de 1960 e 1970, por Mandelbrot (1924-2010). Trata-se de um tipo de matemática para estudar as características dos fenômenos irregulares da natureza, mas teve seu escopo de atuação largamente ampliado e acabou migrando para outras áreas do conhecimento, a exemplo da própria Linguística.

Mandelbrot cunhou o termo fractal do adjetivo latino *fractus*, do verbo *frangere*, que significa *quebrar* e publicou resultados das suas pesquisas no livro intitulado *The Fractal Geometry of Nature* (Mandelbrot, 1977). Em seus estudos, ele explicitou que, se algumas características da natureza são geométricas na perspectiva tradicional, como a lua cheia que se assemelha a um círculo, a maior parte dos fenômenos naturais é irregular (p. ex., as montanhas não são cones nem as nuvens são esferas) e não é contemplada por essa abordagem tradicional da geometria. Um fractal é um objeto que mantém sua forma invariável, quando medido em qualquer escala; é obtido por processos recursivos e apresenta auto-semelhança ou auto-similaridade, complexidade infinita e dimensão<sup>2</sup>.

A sua propriedade de auto-similaridade ou auto-semelhança diz respeito ao fato de os padrões fractais serem repetidamente identificados em escala descendente, de tal maneira que as suas partes serão, em qualquer escala, semelhantes ao seu todo, na forma; entre exemplos existentes para ilustrar essa propriedade, geralmente, recorre-se ao clássico caso da couve-flor. Mas o conceito de fractal contempla tantas outras formas da natureza, como as ramificações dos relâmpagos ou dos vasos sanguíneos que repetem o mesmo padrão inúmeras vezes, deixando patente essa propriedade.

<sup>2</sup> Aqui, a dimensão não será abordada.

Para refletir um pouco mais sobre os fractais, recorro ao exemplo da brócolis romanesco, também, já bastante usado para abordar essa noção matemática. Então, ao se tirar um pedaço dessa verdura, esse pedaço retirado se parece com um pequeno brócolis romanesco e, ao se repetir esse procedimento, o novo resultado encontrado será, da mesma forma, similar a um brócolis desse tipo, mas ainda menor; isto porque cada uma de suas partes se parece com o seu todo, portanto, a forma do brócolis romanesco, em todos os níveis de escala, é semelhante ao próprio brócolis romanesco, em sua totalidade.

A auto-similaridade é resultante do fato de os fractais serem elaborados por iteração<sup>3</sup> que, por sua parte, é a repetição consecutiva de um procedimento. O processo recursivo gerador de fractais é aplicado indefinidamente e quanto maior forem as iterações desse processo, mais detalhes serão apresentados, de tal modo que não se chegará a uma imagem final. Um objeto fractal, com qualquer amplificação, não alcançará uma imagem final, pois poderá conhecer amplificações infinitamente. Por conseguinte, haverá sempre reentrâncias e saliências que serão, ainda, menores.

## 2.1 A metáfora da linguagem como um fractal

Mandelbrot (1977), ao postular a matemática dos fractais, tornou possível pensar sobre a linguagem em termos de fractal. Com isto, no final da década de 1990, por exemplo, o autor Melo e Castro (1997) aponta para a necessidade de haver uma Linguística Fractal e concebe a língua portuguesa como fractal polilinguístico. Quase duas décadas depois, Melo e Castro (2016) ainda ratifica que a noção de fractal pode ser empregada pelas ciências humanas e pela Linguística. E nesse diapasão, afirma:

Quem poderá avaliar e medir rigorosamente todos esses fatos humanos e culturais? Eles escapam, tanto quanto o perímetro de uma ilha ou o volume de uma árvore ou de uma nuvem, a métodos de medição linear e de rigor aritmético. Todos são fenômenos ditos caóticos ou seja, fractais.

Igualmente, quem poderá descrever e “medir”, direta e linearmente, as características da fala dos habitantes de uma região, de uma ilha ou de um país? Ou a sua capacidade inventiva, criadora e transformadora, que a prática linguística em si comporta, diferentemente, da dos falantes da mesma língua, mas de outra região, de outra ilha ou de outro país? (MELO E CASTRO, 2016, p. 6).

---

<sup>3</sup>Vale destacar, porém, como faz Nunes (2006), que nem sempre o processo iterativo gerará um fractal.

O sistema linguageiro, como outros fenômenos da vida, apresenta, portanto, dimensões fractais que são propriedades que se acham em escalas ou estágios diversos; fractais, como aqui já observado, são padrões de formas semelhantes, em consequência, de possuírem auto-similaridade e são independentes da escala de observação. E conforme destaca Paiva (2011), na linguagem, são encontradas operações recursivas que se explicam através da noção de fractais.

Pisemskaya (2007), por sua vez, ao abordar a linguagem sob a perspectiva da Teoria do Caos, e, particularmente, ao enfocá-la por meio da perspectiva dos fractais, observa que, se se proceder à descrição da estrutura sintática do sistema linguageiro, se alcançará uma série de padrões sintáticos idênticos. Além disto, considera que os padrões de organização dos sistemas da linguagem, a exemplo dos padrões do sistema fonológico, são idênticos em todas as escalas. Segundo a citada autora, um ato comunicativo de fala, e, eu acrescento, de escrita, é um exemplo de fractalidade e da sua auto-similaridade.

Rodrigues (2015), por sua parte, considera que, em algumas abordagens, o léxico é visto como excessivamente irregular para ser descrito em termos de padrões. Mas, para a referida pesquisadora, esse sistema da linguagem pode ser compreendido como um fractal em que se verificam iterações de padrões, gerativamente. Para ela, a imagem dos fractais aciona a repetição do uso de padrões, da operação de reconhecê-los e de construí-los. A irregularidade, ela salienta, é uma característica dos fractais e essa irregularidade ocorre tão somente sob a ótica teórico-metodológica que não está preparada para enfrentá-la<sup>4</sup>.

A noção de fractalidade é condizente, da mesma maneira, com os padrões semânticos da linguagem e, também, com a noção de gênero, como discuti anteriormente (ALMEIDA, 2020a, p. 177). Assim sendo, nas

[...] cartas de amor, por exemplo, [...] seus padrões característicos são encontrados repetidamente em escalas descendentes, e, portanto, as suas partes, a qualquer escala, são similares à forma do seu todo. Então, 100.000, 10.000, 1.000, 100, 10 cartas de amor possuem similaridades. Posso afirmar que cada parte do corpus pode ser vista

<sup>4</sup> Rodrigues (2015, p. 11) ressalta que “se fizermos zoom in no léxico, observamos o reconhecimento e a construção de padrões. Na história da morfologia lexicalista, esses padrões começaram por ser identificados como Regras de Formação de Palavras, ou seja, como relações entre bases, operações e produtos. Porém, se fizermos zoom in nessas Regras de Formação de Palavras, verificamos que há subpadrões dentro desses padrões maiores, até descermos a partes cada vez mais pequenas e menos diretamente observáveis, como são aquelas que suportam os constrangimentos e as compatibilidades de anexação base/afixo [...] Cabe, pois, ao linguista a tarefa de descobrir esses detalhes cada vez mais pequenos. Nesta tarefa, não se trata de saber cada vez mais de cada vez menos, porque a análise dessas partes mais pequenas está enquadrada na análise do todo e, sem esse enquadramento no todo, a análise não faz sentido”.

como outra pequena parte desse corpus, no seu todo, isto porque cada carta de amor assemelha-se a todas cartas de amor, já que, na parte, está inscrito o todo; a parte não é somente um todo em si mesmo; a parte é, ao mesmo tempo, uma réplica desse todo; isto pensando no que se concebe como gênero e, no caso particular do exemplo, como carta de amor. Então, se eu estudo, por exemplo, a conceptualização do amor, é certo que uma metáfora como AMOR É VIAGEM, ou uma metonímia como SOFRIMENTO POR AMOR, padrões de organização do amor, poderão se fazer presentes, quer em um pequeno volume de cartas, quer em um grande volume de cartas, quer, ainda, em fragmentos de cartas, sejam poucos ou muitos, isto porque são parte do padrão conceptual do AMOR que se acha instanciado em cartas de amor.

Até aqui ficou demonstrado que a noção de fractal tem sido aplicada a diferentes dimensões da linguagem por diversos pesquisadores e essa compreensão fractal da linguagem é sintetizada por Paiva (2011, p. 4), quando afirma que:

quando pensamos em fractais, pensamos também em operações recursivas como as que encontramos no sistema de linguagem. A língua se organiza em escalas, dos fonemas às palavras, das palavras às orações, das orações às unidades enunciativas, das unidades enunciativas ao discurso que se desdobra em outros discursos em um fluxo infinito.

A linguagem, entendida como um fractal, repete, portanto, o seu padrão de organização em escalas menores. Assim, a auto-similaridade é identificada, quando uma parte da linguagem replica seu todo em uma escala menor e isto ocorre indefinidamente pelo processo iteração que compreende a consecutiva repetição de um procedimento, conforme já pontuado anteriormente.

Com a linguagem concebida como um fractal<sup>5</sup>, o texto, produto da linguagem em ação, pode ser, da mesma maneira, concebido como um fractal, e isto, conseqüentemente, possibilita o entendimento do *corpus*, também, como um fractal. Afinal, o *corpus* pode ser entendido como um conjunto específico de textos, como será exposto na seção a seguir.

---

<sup>5</sup> Vale lembrar, como rememorou um parecerista anônimo, a quem eu agradeço, da existência de diferentes estudos em que a linguagem é analisada computacional e matematicamente como fractal. Há estudos que, conforme ainda recordou o citado parecerista anônimo, antecedem aqueles que empregam a metáfora do fractal como uma maneira de articular aspectos da complexidade dos processos de organização e comunicação linguística (NAJAFI, E.; DAROONEH, A.H (2015); KOHLER, R. (2008); SHANON, B. (1993)). Por fim, cumpre destacar que o percurso aqui adotado é uma opção entre as possíveis rotas acadêmicas existentes para tratar da perspectiva fractal no âmbito dos estudos da linguagem.

### 3 O CORPUS SOB A ÓTICA DA TEORIA DA COMPLEXIDADE

Nesta seção, antes de passar a discutir a noção de *corpus* como fractal, inicialmente, faço um breve apanhado sobre o que se entende por *corpus*, bem como procuro compreendê-lo a partir do arcabouço teórico do Paradigma da Complexidade (Morin, 2009 [1999]). Isto é feito, considerando a discussão que empreendi em outro momento (Almeida, 2020a) e aqui retomada, para que seja possível, na sequência, refletir a respeito das aplicações da amostragem por saturação teórica para a constituição de *corpus* em estudos semânticos na perspectiva da LC/SC.

O conceito de *corpus* tem sido definido por posicionamentos polifônicos, mas é possível afirmar que, no âmbito dos estudos sobre a linguagem, um *corpus* pode ser entendido como uma compilação finita de diversos tamanhos de textos autênticos, isto é, de textos que, de fato, circulem socialmente, criados pela necessidade comunicativa humana, emergentes nos mais variados domínios discursivos, orais ou escritos ou orais-escritos, mono e/ou multissemióticos, transformados em objetos ou em ferramenta de pesquisa, a partir de uma seleção realizada com alguma arbitrariedade, mas com base em uma série de critérios pré-estabelecidos por uma pessoa pesquisadora ou por um grupo de pesquisa.

Ao considerar os princípios da complexidade para pensar o *corpus*, entendo, com base no *princípio sistêmico ou organizacional*, que um *corpus* é mais e menos do que a soma das suas partes; portanto, propriedades e qualidades do todo são suprimidas e novas propriedades e qualidades emergem, quando o *corpus* é constituído; assim sendo, por exemplo, o deslocamento de excertos de uma ação de divórcio — partes de um todo — para, por exemplo, um estudo acadêmico — um outro todo — gera emergências, qualidades e propriedades novas que não se verificavam no todo ação de onde foram coletados os referidos excertos, bem como o *corpus* já não é mais o todo ação de divórcio, de sorte que algumas de suas qualidades e propriedades foram suprimidas quando se tornaram parte de um *corpus*.

Além disto, levando em consideração o *princípio “hologrâmico”*, compreendo que a parte está no todo e o todo está na parte, de modo que, em um *corpus* formado por ações de divórcio, uma ação compartilha os mesmos padrões de organização das demais ações de divórcio, portanto, compartilha, também, os padrões de organização da linguagem atinentes ao gênero em pauta, ocorrendo, tão somente, variação na materialidade linguística das ações, isto é, na sua estrutura, conforme a terminologia do Paradigma da Complexidade.



Ademais, para discutir a noção de *corpus*, penso ser necessário recorrer ao *princípio de retroatividade*, a fim de observar que a linearidade na geração do conhecimento é uma ilusão; saberes retroagem sobre saberes anteriores e são repensados em novos contextos. Como a causa age sobre o efeito e esse sobre aquela, no âmbito de uma pesquisa, a geração do *corpus* age sobre os seus resultados, bem como, no plano geral do fazer científico, resultados de pesquisas agem sobre a constituição de *corpus*.

Além disto, o princípio do *circuito recursivo* precisa ser recuperado para que seja possível ponderar sobre o fato de os resultados das investigações sobre a linguagem serem produtos de um sistema científico que possui uma tradição de constituir *corpus* e que essa tradição não se reproduzirá com seus resultados, se *corpora* não forem constituídos pelo sistema científico; assim, os produtos elaborados através de *corpus* são causadores da produção do processo científico e, também, em última instância, de *corpus*.

Ao pensar o *corpus* através da ótica da complexidade, é importante retomar, igualmente, o *princípio dialógico* e considerar que um *corpus* é constituído, por uma desordem que leva a uma nova ordem, isto porque, por exemplo, uma ação de divórcio, gerada por necessidades comunicativas próprias do domínio discursivo jurídico, portanto, em estado de ordem nesse âmbito, conhece uma desordem, quando ocorre o seu deslocamento ou ainda a sua fragmentação, para se tornar parte de um *corpus*; contudo essa desordem está no âmago de um novo estado de ordem, uma vez que, enquanto *corpus*, esse texto ou seus fragmentos farão parte de um novo texto de outro gênero (uma Dissertação de Mestrado, p. ex.), sendo, então, parte de um outro domínio discursivo, no caso, do domínio discursivo científico. Logo, o novo texto, constituído também por meio de um *corpus*, será gerado através da dialógica ordem-desordem-ordem.

É necessário atentar, da mesma maneira, para o *princípio da reintrodução do conhecimento em todo conhecimento* e, com isto, compreender que, para gerar um *corpus* com a finalidade de produzir saberes sobre a linguagem, ocorrem (re)construções empreendidas por uma pessoa pesquisadora que, inserida em um espaço científico geo-sócio-histórico-cultural-político-ideológico, conceptualiza-categoriza, para assim gerar os saberes pretendidos.

Por fim, o trabalho de constituição de *corpus* é ainda sensível às condições iniciais, de tal sorte que, uma lacuna, aparentemente insignificante, em conjunção com outras lacunas, em tese, desimportantes, podem gerar diferenças significativas na compreensão de um dado fenômeno da linguagem; por exemplo, a negativa a uma resposta para um

inquérito léxico-semântico de natureza dialetológica, somada a outras pequenas respostas negativas, pode ter implicações consideráveis na constituição do *corpus* e nos resultados alcançados com a realização de uma pesquisa elaborada com esse mesmo *corpus*.

Feitas essas breves considerações a respeito do *corpus* sob a ótica do Paradigma da Complexidade, passo, na sequência, a refletir brevemente sobre o *corpus* entendido como um fractal.

### 3.1 O *corpus* como fractal

Pensar o *corpus* como fractal sustenta teoricamente a mudança do enfoque do fazer científico acerca da linguagem do quantitativo para o qualitativo, da estrutura para o padrão de organização, uma vez que, ao compreendê-lo desta maneira, se torna possível focar os padrões da linguagem, em consequência, os padrões do fenômeno linguageiro a que se propõe estudar; assim, vai-se além da estrutura, ou seja, da matéria linguística, isto porque, enquanto essa cinge quantidades, aqueles cingem qualidades, conforme já assinalou Capra (2006 [1996]) e também Capra e Luisi (2014).

Como já visto, cada pequena seção de um fractal pode ser vista como uma réplica em tamanho menor do seu todo. Por conseguinte, em um *corpus* constituído por ações de divórcio, o texto de uma ação de divórcio é semelhante a todos os outros textos desse gênero, isto, porque a parte está no todo e o todo está na parte, logo, as ações de divórcio ou até mesmo os seus fragmentos exibem, como já também destacado, o padrão de organização desse gênero, por apresentarem auto-similaridade.

Além disso, é preciso considerar que, por ser a ação de divórcio replicável em inúmeras ações de divórcio, não será possível alcançar a uma imagem final de todas as ações de divórcio em nenhum *corpus*, pois são incontáveis as possibilidades de sua amplificação. Portanto, nunca se alcançará a pretendida representatividade quantitativa do *corpus*<sup>6</sup>, pois sempre se estará em defasagem em relação aos textos que iteraram o padrão de organização de um determinado gênero e de sua linguagem.

---

<sup>6</sup> Silva e Silva (2013) observam que pesquisadores ponderam que a representatividade de um *corpus* se atrela ao seu tamanho, de tal sorte que consideram que esse será mais representativo da linguagem, quanto maior for o seu conjunto textual. E embora sejam identificadas tentativas de determinar quantidades indicativas de tamanhos, indo do *corpus* pequeno ao grande, o estabelecimento de sua representatividade medida por seu tamanho não forma consenso entre os pesquisadores. Neste sentido, Charaudeau (2011, p.1), inclusive, destaca que “[...] a hipótese da exaustividade — velho sonho da atitude positivista — não se sustenta mais, e isso, apesar do desenvolvimento da chamada linguística de *corpus*”.

Mas, por outro lado, esses padrões podem ser acessados em pequenas parcelas de textos, uma vez que a própria linguagem e os textos resultantes do seu uso nas interações humanas são fractais.

Após ter retomado a noção de *corpus* a partir da metáfora dos fractais, a seguir, passo a pensar sobre o uso da Técnica da Saturação no âmbito da elaboração de *corpus* em estudos semânticos, abordando, em particular, o fenômeno da conceptualização a partir da LC/SC.

#### 4 A TÉCNICA DA SATURAÇÃO E A CONSTITUIÇÃO DO CORPUS

Se é possível acessar os padrões de organização da linguagem através de pequenas porções de material textual, como aqui já visto anteriormente, como ter segurança em relação à representatividade qualitativa do *corpus*? Essa questão sempre se impõe para quem se dedica aos estudos da linguagem em perspectiva qualitativa.

Além disto, como a linguagem possui a propriedade de iteração, como também já observado, jamais será possível reunir toda produção linguageira que testemunhe um fenômeno linguístico. Então, o fechamento do *corpus* por exaustão é inalcançável, pois, diante da complexidade infinita, da capacidade de amplificação da linguagem, se estaria sempre a um passo atrás de uma nova realização desse fenômeno. Logo, se a delimitação do *corpus* por exaustão é, posso afirmar, uma quimera, ainda que seja possível constituir *corpora* de proporções inimagináveis antes dos últimos avanços tecnológicos, outras perguntas se colocam e dizem respeito ao como é possível interromper o levantamento de material textual quando da constituição de *corpus* e qual seria a justificativa para essa interrupção.

A amostragem por saturação é uma ferramenta conceitual viável para que a pessoa pesquisadora possa fechar o tamanho final do *corpus*, de maneira a fornecer uma base sólida para a realização de estudos qualitativos sobre a linguagem, pois, apesar de, em pesquisas alicerçadas a partir de métodos compreensivo-interpretativos, a quantidade ser secundariamente importante, quando comparada às dimensões contextuais, é preciso delimitar o fechamento da coleta de dados, assim como é fundamental que se tenha segurança para fazê-lo.

Em pesquisas de natureza qualitativa, o processo de amostragem por saturação teórica é um modo seguro para estabelecer esse momento de interrupção da coleta de dados, porque oferece o rigor metodológico necessário à investigação científica, de sorte que, com transparência, possibilita demonstrar como se deu a saturação dos dados.

Quando fica constatado que não se depreendem mais, no campo de observação, elementos novos para embasar a discussão acerca do

fenômeno pesquisado, ou melhor, quando a interação entre a pessoa pesquisadora e os textos já não fornece mais esses novos elementos para o balizamento e para o aprofundamento da compreensão daquilo que se estuda, ocorre o ponto de saturação e é o princípio auto-similaridade da linguagem, aqui antes abordado, que justifica o uso dessa perspectiva para a constituição do *corpus*.

Outra questão que se apresenta a quem pretende empregar a técnica da saturação teórica diz respeito ao como proceder para aplicá-la. Um caminho possível é primeiramente delimitar o fenômeno linguageiro a ser abordado e o(s) domínio(s) discursivo(s) a partir do(s) qual(is) os textos constituintes do *corpus* serão coletados; depois, interagir com esses textos e proceder à leitura na íntegra de cada um deles, a fim de identificar a(s) categoria(s) a serem enfocadas.

Com isto, o estudo do fenômeno abordado vai sendo realizado concomitantemente à própria constituição do *corpus* e, para visualizar os dados encontrados, geralmente, é produzido um quadro em que a(s) categoria(s) levantada(s) é(são) inserida(s) com a expressão da linguagem — mono e/ou multissemiótica — que permitiu delimitá-las, sendo que esse quadro possibilita identificar a regularidade das categorias. Com repetição desse procedimento, será possível chegar ao ponto em que a saturação ocorre, isto é, quando já começam as redundâncias ou repetições, quando a inclusão de informações já não modifica o entendimento do que se busca compreender e, assim, é possível determinar a validade do conjunto de dados coletado.

O trabalho de aplicação da amostragem por saturação teórica não é automático, de sorte que o engenho de quem pesquisa delimita quando interromper a captação de novos elementos, determina quando não é mais produtivo investir na coleta de dados, porque a persistência em coletar novos componentes não trará contribuições significativas para aperfeiçoar, com base nos dados, as reflexões acerca do fenômeno destacado. Assim sendo, o uso dessa técnica coaduna com o princípio da reintrodução do conhecimento em todo conhecimento, conforme já exposto na segunda seção deste artigo (*O corpus sob a ótica da Teoria da Complexidade*).

As propriedades, as determinações e os conhecimentos atinentes aos objetos de pesquisa são, como destaca Fontanella, Ricas e Turato (2007), organizados sociocultural e psicologicamente, de tal modo que não é possível desprezar o fato de os dados serem elaborados através de conceitos prévios daquele que pesquisa não sendo assim passivamente advindos da realidade. Os fenômenos da linguagem existem independentemente do linguista, mas só se transformam em objetos de estudo da Linguística, quando um observador(a)-pesquisador(a) os transforma em objeto de

pesquisa, quando maneja um arcabouço teórico-metodológico científico, elaborado com base na imaginação da sua mente científica corporificada.

Enfim, como destacam os mesmos autores,

na medida em que as ferramentas utilizadas na constatação de saturação [...] são de ordem [...] cognitiva (envolvendo a percepção do pesquisador e seu domínio teórico), vários parecem ser os fatores que intervêm nesse processo, podendo influenciar a decisão de interromper o recrutamento da amostra por considerá-la saturada, usando agora a expressão no sentido de "amostra completa". Cabe ressaltar, ainda, que qualquer projeto de pesquisa se insere num contexto que extrapola questões puramente metodológicas, situando-se num determinado contexto político de produção científica. Os objetivos de pesquisa inserem-se também neste contexto mais amplo e, como vimos, eles influenciam diretamente todos os aspectos do processo de amostragem. (FONTANELLA; RICAS; TURATO, 2007, p. 5).

Nada disto tem qualquer relação com práticas que indiquem o abandono do rigor e da transparência metodológicos, pois sempre é necessário haver o compromisso com a ciência. Portanto, quando se aplica qualquer técnica científica, incluindo, conseqüentemente, a amostragem por saturação, faz-se necessário explicitar, de modo inteligível aos pares, os critérios adotados para sua utilização.

A metodologia com seus procedimentos não é uma receita predeterminada que só cabe a quem pesquisa segui-la, mas isto não exige a pessoa pesquisadora de descrever, com rigor e transparência, os procedimentos metodológicos, a exemplo do processo empregado para identificar a saturação, isto é, quando proceder à interrupção da coleta de dados, isto porque esse comportamento rigoroso e transparente atende ao critério de confiabilidade das pesquisas científicas de natureza qualitativa.

Como observaram Fontanella, Ricas e Turato (2007), apesar de nenhum discurso ser igual a outro, como inclusive apontam os estudos da complexidade, todos, como também já comentado em outra seção deste artigo, possuem o mesmo padrão de organização (Maturana; Varela, 2001 [1984]), por isto, necessariamente, expressarão elementos em comum, contudo, por serem elaborados por meio de perspectivas distintas, vão também apresentar diferenças na estrutura, ou seja, vão conter variação em sua materialidade (Almeida, 2020b).

Na constituição de um *corpus*, quando os primeiros textos sobre um determinado tema são acessados, serão identificadas mais diferenças entre eles, entretanto, conforme se vai tendo contato com outros textos, essa diversidade vai diminuindo até se tornar irrelevante, porque os padrões se repetem e porque as perspectivas possíveis não são infinitas, já que são

definidas pela corporeidade da espécie humana, conforme explicitam Lakoff e Johnson (1999).

Além disso, a diversidade vai diminuir e tornar-se irrelevante, porque pode fugir aos objetivos estabelecidos para a pesquisa e pelo fato de o grau de relevância de um achado atrelar-se sempre a tais objetivos; assim sendo, um dado, considerado pouco relevante em face de um objetivo, pode ser categorizado como basilar para outro. Isto demonstra que há espaço para a inserção de quem pesquisa como elemento determinante nos resultados de uma pesquisa. Logo, caberá à pessoa pesquisadora, quando um conjunto de percepções a partir da amostra se repete, definir se as diferenças apresentadas podem ou não ser mais valorizadas e exploradas.

Então, diante dos objetivos de quem pesquisa, em face de distintos tipos de informações, a saturação acontecerá em momentos diversos. Assim, são os objetos e os objetivos da pessoa pesquisadora que determinarão o momento da saturação (Fontanella; Ricas; Turato, 2008). Ademais, a saturação das categorias não ocorre necessariamente no mesmo texto; por conseguinte, uma categoria pode saturar em um texto diferente de outra.

O número de recorrências por si só não é fundamental para a constituição da amostragem por saturação, uma vez que demonstra, tão somente, o que seria mais ou menos comumente falado e/ou escrito e expressará, assim, convencionalidades e particularidades, pensando, respectivamente, em termos de mais ou menos ocorrências. Ao promover uma abordagem qualitativa do estudo científico da linguagem, a pessoa pesquisadora poderá debater o contínuo dos padrões de organização do fenômeno por ela estudado, discutindo perspectivas, especificidades, generalizações expressas nesse contínuo. Mas, para que seja possível determinar, com segurança, o final da coleta do material textual constituinte do *corpus*, é sugerido que, depois de identificada a saturação, seja acrescida a leitura de mais  $1/3^7$  de textos.

Se a saturação de uma dada categoria ocorre, por exemplo, no décimo texto, precisam ser lidos mais 3 textos para ratificar o resultado da saturação (Nascimento *et. al*, 2018); isto demonstra que o pensamento matemático não é abandonado, mas deixa de ser regido pela doxa da quantificação, como discute Maffesoli (2016), e passa a ser requerido a partir da razão imaginativa, tal como pensada no seio da LC/SC.

#### 4.1 A Técnica da Saturação em LC/SC

<sup>7</sup> A leitura de mais  $1/3$ , em vez de qualquer outra fração, é sustentada academicamente pelo pensamento expresso por Abreu (2020) acerca da divisão triádica, baseada no corpo humano.

Como destaca Santana (2019, p. 83):

a Técnica da Saturação Teórica é muito utilizada, na coleta de dados, em pesquisas nas áreas de Administração (Falqueto; Farias, 2016; Thiry-Cherques, 2009) e Saúde (Fontanella et. al., 2011; Fontanella; Ricas; Turato, 2008; Turato, 2005). A forma de utilização mais comum do critério da saturação é o da aplicação de entrevistas, mas também são usadas nas técnicas de observação, grupo focal e aplicação de questionários.

Em LC/SC, a amostragem por saturação vem sendo paulatinamente adotada e a tese de Santana (2019) oferece um exemplo de adoção dessa forma de constituir *corpus* no âmbito dos estudos da linguagem, inclusive, partindo de textos escritos no passado, uma vez que a referida autora constituiu seu *corpus* com cartas de amor escritas entre os séculos XIX e XX. Também, a tese de Barreto (2021) traz o uso dessa amostragem na constituição do *corpus*, mas diferentemente de Santana (2019) que trabalhou com textos de um único gênero, Barreto (2021) utilizou textos do domínio discursivo jornalístico pertencentes a diversos gêneros, publicados desde os Oitocentos até a atualidade, portanto o *corpus* da sua tese contemplou textos do presente, além daqueles produzidos no passado, como os que foram usados no trabalho de Santana (2019).

Ademais, Duque (2020) empregou a mesma abordagem para formar um *corpus* verbo-imagético, constituído por charges, para o estudo que desenvolveu sobre os sentidos da COVID. Assim, é possível constatar que, em Linguística, a amostragem por saturação pode ser utilizada tanto para a constituição de *corpus* a partir de textos do passado quanto do presente, sejam esses verbais e/ou multissemióticos, indo, portanto, além dos gêneros entrevista e daqueles utilizados em outros campos do conhecimento científico que já empregam com mais recorrência essa técnica, conforme assinalado na citação de Santana (2019) que abre essa subseção deste artigo.

#### **4.1.1 Exemplos de aplicações da Técnica da Saturação em estudos sobre a conceptualização**

A seguir, para exemplificar como a referida amostragem por saturação tem sido aplicada em estudos produzidos no âmbito da LC/SC, para tratar do fenômeno da conceptualização, será sintetizado o *modus*

*operandi* de Santana (2019) e de Duque (2020)<sup>8</sup> para constituírem respectivamente o *corpus* desses seus estudos, aqui antes citados.

Para identificar o ponto de saturação, nas cartas de amor que teve acesso para a constituição do *corpus* de sua tese, Santana (2019) procedeu à leitura individual das cartas de cada escrevente, seguindo a sequência cronológica da escritura da correspondência, e, com a leitura de cada texto, foi localizando as expressões linguísticas que materializavam as conceptualizações do AMOR. Conforme selecionava as expressões, também, identificava o domínio-fonte requerido para a conceptualização em cada expressão e ia colocando essas informações em um quadro. Para exemplificar o comportamento adotado pela referida pesquisadora, na sequência, será exposta uma adaptação dos quadros constantes da sua tese que se referem ao material textual coletado das cartas do conceptualizador-escrevente Jayme.

Quadro 01 – Identificação dos domínios no corpus de Santana (2019)

<b>INFORMAÇÕES DO QUADRO 2</b>	<b>INFORMAÇÕES DO QUADRO 3</b>	<b>INFORMAÇÕES DO QUADRO 4</b>
Carta 1 – 12/01/1937	Carta 2 – 19/01/1937	Carta 3 – 24/01/1937
Domínios identificados	Recorrência de domínio	Recorrência de domínio
OBJETO POSSUÍDO APROXIMAÇÃO	OBJETO POSSUÍDO -----	OBJETO POSSUÍDO APROXIMAÇÃO SOFRIMENTO CORAÇÃO
	Novos domínios	Novos domínios
	SER HUMANO SOFRIMENTO CORAÇÃO	-----

Fonte: adaptado de Santana (2019)

Na primeira coluna do quadro 01, antes apresentado, aparecem expostos os dados constantes do quadro 2 da tese de Santana (2019), no qual foram registrados os domínios identificados nas expressões linguísticas que instanciam as conceptualizações de amor na primeira carta de Jayme. No terceiro quadro da referida tese, aduzido aqui na segunda coluna do quadro 01, foram documentados os domínios recorrentes, portanto já localizados na primeira carta, assim como os novos domínios instanciados na segunda carta do mesmo conceptualizador-escrevente. Por fim, no quadro número

<sup>8</sup> Esses autores foram selecionados, em detrimento de outros, pois foram os primeiros a demonstrarem a validade da proposta de Almeida (2020a), desenvolvida em 2018, mas publicada em 2020; assim, Santana (2019) é a primeira a empregar essa proposta para gerar corpus verbal e Duque (2020) é o primeiro pesquisador a utilizá-la para constituir corpus multissemiótico, especificamente, verbo-imagético.



4, transcrito na terceira coluna do quadro 01, foi adotado o mesmo comportamento empregado para o preenchimento do terceiro quadro da referida tese, de modo a apresentar as recorrências e os novos domínios.

Além disto, em uma tabela, Santana (2019) reproduziu todo procedimento realizado para cada uma das cartas que leu do conceptualizador-escrevente Jayme, promovendo a visualização dos domínios identificados na correspondência dele:

Tabela 1 – Ponto de saturação nas cartas do escrevente Jayme

<b>Cartas Domínios</b>	1	2	3	4	5	6	7	8	9	10	11	12	13	14
OBJETO POSSUÍDO	<b>X</b>	X	x	x	x	x	x			x		x		
APROXIMAÇÃO	<b>X</b>		x	x	x		x	X	x	x	X			
SER HUMANO		<b>X</b>				x							x	
SOFRIMENTO		<b>X</b>	x	x		x	x		x	x				x
CORAÇÃO		<b>X</b>	x	x	x	x							x	
PAIXÃO					<b>X</b>					x				
GOSTAR						<b>X</b>							x	
DESEJO								<b>X</b>			X	x		x
SAUDADE										<b>X</b>				

Fonte: Santana (2019, p. 85)

Na primeira coluna da tabela 01, foram inseridos os domínios constantes das cartas do conceptualizador-escrevente Jayme. Na primeira linha, as cartas escritas pelo referido conceptualizador-escrevente foram numeradas de 1 a 14; essas quatorze cartas que foram utilizadas para a constituição do *corpus* foram organizadas em ordem cronológica; assim sendo, o número 1, constante da primeira linha da tabela, refere-se à primeira carta lida, enquanto o número 14 diz respeito à última carta lida para confirmar o ponto de saturação. Nessa tabela, o **X**, maiúsculo e em negrito, que já aparece a partir da segunda linha, foi usado para indicar a presença de um novo domínio e o x, minúsculo sem negrito, foi empregado para informar os domínios que se repetem, ou seja, os domínios recorrentes.

Conforme Santana (2019), o ponto de saturação, na correspondência de Jayme, deu-se na carta número 11 que aparece com sombreamento cinza na tabela 01 antes exposta neste artigo. Assim sendo, a partir daí, não foram mais localizadas novas informações que pudessem contribuir para o estudo que foi empreendido. Com a detecção do ponto de saturação na décima primeira carta, foi ainda realizada a leitura de mais três cartas, totalizando 14 cartas; isto foi feito para confirmar se havia, de fato, ocorrido o ponto de saturação.

A tabela 01, trasladada da tese de Santana (2019, p. 85), possibilita verificar que os acréscimos de novos domínios ocorreram mais nas primeiras cartas. Logo,

a carta 2, por exemplo, confirma o domínio do OBJETO POSSUÍDO, da carta 1, e acrescenta SER HUMANO, SOFRIMENTO e CORAÇÃO. Posteriormente, os acréscimos ou vão diminuindo ou não existem até que deixam de aparecer, a partir da carta 11. Após a leitura de mais três cartas, confirmamos a repetição. Desse modo, para esse escrevente, consideramos todas as ocorrências das 14 (catorze) primeiras cartas. Isso não significa que, nas cartas de todos os escreventes, o ponto de saturação ocorrerá na 11ª carta, ao contrário, pode ser na 4ª, 6ª, 13ª [...]

Duque (2020), por sua parte, elaborou e delimitou previamente um *corpus* constituído por charges, portanto, por textos verbo-imagéticos, produzidos na atualidade, diferentemente de Santana (2019) que trabalhou com textos verbais, produzidos no passado. Para a constituição desse *corpus*, o referido pesquisador contou com o auxílio da ferramenta de busca do Google Imagens e empreendeu sua busca, nessa ferramenta, através da inclusão das seguintes expressões linguísticas: pandemia OR COVID-19 OR coronavírus AND charge. Com este procedimento, teve acesso a 613 imagens. Com o estudo dessas imagens, foi possível que excluísse cartuns, tirinhas, quadrinhos e fotografias, restando 515 charges. De posse desses verbo-imagéticos, o referido autor aplicou a Técnica da Saturação Teórica. O ponto de saturação foi alcançado com 60 textos, de modo que os outros verbo-imagéticos acessados não lhe forneceram dados relevantes para o seu estudo, pois repetiram os padrões de organização semânticos.

Duque (2020), inspirado por Falqueto e Farias (2016), propôs 5 etapas para a identificação do ponto saturação da amostra de charges:

- a) Delimitação dos frames subjacentes às pistas de linguagem nas charges, sendo que as categorias foram sendo definidas com o estudo de cada charge alcançada pela pesquisa feita ao Google;
- b) Definição do roteiro de pesquisa, a saber: 2.1) Como o Coronavírus é representado graficamente? 2.2) Que personagens estão na charge? 2.3) Qual o cenário em que ocorre o evento da charge? 2.4) Qual é o evento da charge? Esse roteiro colaborou para a identificação do ponto de saturação;
- c) Levantamento das informações novas x informações já confirmadas; com isto, foi possível proceder à identificação das informações relevantes e categorizá-las. Com a exploração de cada charge, antes de seguir para o próximo texto, informações novas foram diferenciadas das outras localizadas nos textos anteriormente consultados. Não foram identificados frames novos que se mostrassem relevantes para o estudo a partir da charge número 41. Apesar disto, foram, ainda, analisadas outras charges com a finalidade de confirmar o ponto de saturação. No conjunto textual estudado, foram localizados 6 frames recorrentes: GOVERNO; NEGACIONISMO; PROTEÇÃO; CONTAMINAÇÃO; AGLOMERAÇÃO e ECONOMIA;

- d) Registro do que foi encontrado de novo em cada coleta em uma tabela; esse procedimento tornou possível uma melhor visualização da saturação dos frames acionados no âmbito da conceptualização em pauta, como a seguir apresentado:

Tabela 02 – Saturação teórica das charges

Categorias	Charges																			
	1	2	3	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
GOVERNO	-	1	1	1	-	-	1	1	-	-	-	-	-	1	-	-	1	-	-	-
PROTEÇÃO	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NEGACIONISMO	-	-	-	-	-	-	1	-	1	1	1	-	-	1	1	-	-	-	1	0
ECONOMIA	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	1	-	1	-	-
CONTAMINAÇÃO	1	-	-	-	-	1	-	-	-	-	-	-	1	-	-	-	-	-	-	-
AGLOMERAÇÃO	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Fonte: Duque (2020, p.119)

Na tabela aqui antes transcrita, “nas linhas, foram atribuídos o valor de 1, para informar que havia, pelo menos, uma nova informação e -, para indicar que não foi encontrada nenhuma nova informação na respectiva charge” (Duque, 2020, p. 119).

- e) Confirmação da saturação de cada frame. Na referida tabela, é possível constatar que o frame AGLOMERAÇÃO apresentou saturação na charge 26. Apesar da identificação da saturação, ainda, foram analisadas outras charges, pois o ponto de saturação de outros frames não havia ainda sido alcançado. Como já informado, a saturação geral ocorreu na charge 41, mas 19 novas charges foram acessadas, totalizando 60 textos verbo-imagéticos, para, com isto, obter maior segurança em relação à delimitação do ponto de saturação.

Os estudos de Santana (2019) e de Duque (2020) confirmam a validade da amostragem por saturação teórica que, por sua vez, possibilita refletir sobre padrões de organização semântica da linguagem. Dito isto, a seguir, passo às últimas palavras deste artigo.

## 5 CONSIDERAÇÕES FINAIS

A noção de fractais, utilizada para amparar as discussões aqui propostas, é sustentada pela razão sensível da TC e emerge das vivências experienciadas por todos — cientistas e não cientistas — e, por isto mesmo, é traduzida, não só pela ciência, mas também, como revela a epígrafe deste texto, pela sensibilidade razoável de poetas, como Fernando Pessoa e seu heterônimo Álvaro de Campos.

No âmbito da ciência, domínio discursivo que enseja o artigo que ora se finda, a concepção científica da linguagem em termos de fractal sustenta o emprego da técnica da saturação teórica para a constituição de

*corpus* que, por ser uma amostra da linguagem, é também compreendido em termos fractais.

Os estudos de Santana (2019) e de Duque (2020), entre outros, como os trabalhos de Almeida (2020b) e de Barreto (2021), comprovam a eficácia dessa técnica para a geração de *corpus* em pesquisas de natureza qualitativa empreendidas em Linguística, em particular, em LC/SC.

Isto posto, antes de concluir efetivamente este texto, retomo a discussão aqui antes produzida no sentido de ratificar que compreender fenômenos da linguagem a partir de uma abordagem qualitativa do fazer científico não significa adotar uma postura sem rigor e transparência, não quer dizer que sejam desconsiderados os aspectos e os procedimentos metodológicos, sob argumento simplista de serem esses aspectos e procedimentos meras preocupações do positivismo.

E ainda, levando em consideração a importância do labor metodológico para os estudos da linguagem e entendendo-o como parte vital do fazer científico, mas sem ser, por isto, sua camisa de força, neste artigo, reafirmo, foi colocado em pauta o como proceder para estabelecer o tamanho final da amostra textual a ser estudada, sugerindo, para isto, o emprego da técnica da saturação que, por sua parte, contém procedimentos que, como aqui visto, podem ser adaptados às particularidades de cada pesquisa, demonstrando o protagonismo da pessoa pesquisadora na definição do corpus e, conseqüentemente, no alcance dos resultados de suas pesquisas.

## REFERÊNCIAS

- ABREU, A. S. **Figuras de construção e imagens: aspectos cognitivos e funcionalidade retórica**. YouTube, 23 nov. 2020. Disponível em: [ICIGELP - DIA 25: Antonio Suarez Abreu - YouTube](#) Acesso em: 01 jun. 2022.
- ALMEIDA, A. A. D. A tessitura do conhecimento: o corpus na construção de estudos semânticos sócio-histórico-cognitivos. In: SOUZA, R.; BORGES, R.; Almeida, I.; SOUZA, D. **Filologia em diálogo: descentramentos culturais e epistemológicos**. Salvador: Arte& Memória, 2020a, p. 148-184. Disponível em: [https://www.memoriaarte.com.br/\\_files/ugd/33823c\\_484db24267204deab916bd4bd15eedf0.pdf](https://www.memoriaarte.com.br/_files/ugd/33823c_484db24267204deab916bd4bd15eedf0.pdf) Acesso em: 10 out. 2021.
- ALMEIDA, A. A. D. Estamos sempre em guerra? Estudo cognitivo sócio-histórico de uma metáfora da Gripe Espanhola e da COVID-19. **Estudos Linguísticos e Literários**, Salvador, v. 69, p. 366-395. 2020b. Disponível em: <https://periodicos.ufba.br/index.php/estudos/article/view/44310/24478> Acesso em: 05 out. 2021.
- ALMEIDA JÚNIOR, A. T. **Espaços e atratores – estratégias de categorização na emergência de inferências sobre a conceitualização de violência**. 2013. Tese (Doutorado em Linguística) – Centro de Humanidades, Universidade Federal do Ceará, Fortaleza, 2013.
- BARRETO, D. P. **Conceptualização/categorização do estupro: um estudo sócio-histórico-cognitivo**. 2021. Tese (Doutorado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2021.
- BERNÁRDEZ, E. De la "lingüística catastrofista" a la lingüística cognitiva. **Revista de Filología Alemana**, Madrid, n.2, p. 181-199, 1994. Disponível em: [\(PDF\) De la "lingüística catastrofista" a la lingüística cognitiva \(researchgate.net\)](#) Acesso em: 05 out. 2021.
- BRANDT, P. A. *La Charpente modale du sens*, Amsterdam: John Benjamins, 1992.
- CAPRA, F. **A teia da vida: uma nova compreensão científica dos sistemas vivos**. Trad. Newton Roberval Eicheberg. São Paulo: Cultrix, 2006 [1996].
- CAPRA, F.; LUISI, P. L. **A visão sistêmica da vida: uma concepção unificada e suas implicações filosóficas, políticas, sociais e econômicas**. Trad. M. T. Eicheberg e N. R. Eicheberg. São Paulo: Cultrix, 2014.
- CHARAUDEAU, P. Dize-me qual é teu corpus, eu te direi qual é a tua problemática. **Diadorim**, Rio de Janeiro, v. 10, p. 01-23, 2011. Disponível em: <http://www.revistadiadorim.lettras.ufrj.br>. Acesso em: 05 jun. 2022.
- DUQUE, P. H. *A covid-19 em charges: uma análise baseada em frames*. **Estudos Linguísticos e Literários**, Salvador, v. 69, p. 106-127, 2020. Disponível em: <https://periodicos.ufba.br/index.php/estudos/article/view/44290> Acesso em: 04 mai 2022.

FERNÁNDEZ JAÉN, J. Semántica histórica y teoría del caos. **Res Diachronicae**, [s.l.], v. 6, p. 21-39, 2008. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=2718809> Acesso em: 04 mai. 2022.

FONTANELLA, B. J. B; RICAS, J. TURATO, E. R. Amostragem por saturação em pesquisas qualitativas em saúde: contribuições teóricas. **Cad. Saúde Pública**, Rio de Janeiro, v. 24, n. 1, p. 17-27, 2008. Disponível em: <https://www.scielo.org/article/csp/2008.v24n1/17-27/> Acesso em: 04 out.2021.

KOHLER, R. Are there Fractal Structures in Language? Units of Measurement and Dimensions. **Linguistics. Journal of Quantitative Linguistics**, v. 4, p.122-125, 1997. Disponível em: <https://www.tandfonline.com/action/showCitFormats?doi=10.1080%2F09296179708590085> Acesso em: 05 out. 2022.

LAKOFF, G.; JOHNSON, M. **Philosophy in the flesh**: the embodied mind and its challenge to western thought. New York: Basic Books, 1999.

MAFFESOLI, M. **A ordem das coisas**: pensar a pós-modernidade. Trad. A. CHIQUIERI Rio de Janeiro: Forense, 2016.

MANDELBROT, B. B. **The Fractal Geometry of Nature**. New York: W. H. Freeman, 1977.

MATURANA, H; VARELA, F. **A árvore do conhecimento**: as bases biológicas da compreensão humana. Trad. Humberto Mariotti e Lia Diskin. São Paulo: Palas Athena, 2001 [1984].

MELO E CASTRO, E. M. As ilhas do arquipélago. **Via Atlântica**, [s.l.], v. 1, n. 1, p. 108-115, 1997. Disponível em: <https://www.revistas.usp.br/viaatlantica/article/view/48675> Acesso em: 04 out. 2021.

MELO E CASTRO, E. M. Uma concepção fractal da língua portuguesa. Comunicação inaugural. In: CONGRESSO INTERNACIONAL PELOS MARES DA LÍNGUA PORTUGUESA, 3, 2016, Aveiro. **Anais** [...] Aveiro: Departamento de línguas e culturas da Universidade de Aveiro, mai. 2016. Disponível em: <http://www.musarara.com.br/uma-concepcao-fractal-da-lingua-portuguesa>. Acesso em: 04 out.2021.

MORIN, E. **A cabeça bem-feita**: repensar a reforma, reformar o pensamento. Trad. Eloá Jacobina. Rio de Janeiro: Bertrand, 2009 [1999].

NAJAFI, E, DAROONEH, A. H. The Fractal Patterns of Words in a Text: A Method for Automatic Keyword Extraction. **PLoS One**, v. 10, n. 6, p. 1-18, jun., 2015. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0130617> Acesso em: 05 out. 2022.

NASCIMENTO et al. Saturação teórica em pesquisa qualitativa: relato de experiência na entrevista com escolares. **Revista Brasileira de Enfermagem**, [s.l.], v. 71, n. 1, p. 243-248, 2018. Disponível em: <http://old.scielo.br/scielo.php?pid=S0034->

[71672018000100228&script=sci\\_arttext&tIng=pt#:~:text=Quando%20o%20roteiro%20de%20entrevista,dever%C3%A3o%20ser%20realizadas\(2\)](#). Acesso em: 04 out.2021.

PAIVA, V. L. M. de O. A metonímia como processo fractal multimodal. **Veredas**, Juiz de Fora, v.14, n.1, p. 07-19, 2010. Disponível em: <http://www.ufjf.br/revistaveredas/files/2010/04/ARTIGO-1.pdf>. Acesso em: 05 fev. 2022.

PAIVA, V. L. M. de O. Língua(gem) como sistema complexo e multimodalidade. **ReVEL**, [s.l.], v. 14, n. 26, p. 331-344, 2016. Disponível em: [www.revel.inf.br](http://www.revel.inf.br) Acesso em: 04 out. 2021.

PAIVA, V. L. M. de O. Linguagem e aquisição de segunda língua e na perspectiva dos sistemas complexos. In: BURGO, V. H.; FERREIA, E.F.; STORTO, L. J. **Análise de textos falados e escritos: aplicando teorias**. Curitiba: CRV, 2011. p. 71-86.

PAIVA, V. L. M. O. Modelo fractal de aquisição de línguas. In: BRUNO, F. C. (Org.). **Reflexão e prática em ensino/aprendizagem de língua estrangeira**. São Paulo: Clara Luz, 2005, p. 23-36. Disponível em: <http://www.veramenezes.com/modelo.htm>. Acesso em: 04 out.2021.

PETITOL, J. Syntaxe topologique el grammaire cognitive. **Langages**, [s.l.], v. 25, n. 103, 1991, p. 97-128. Disponível em: [https://www.persee.fr/doc/lgge\\_0458-726x\\_1991\\_num\\_25\\_103\\_1610](https://www.persee.fr/doc/lgge_0458-726x_1991_num_25_103_1610). Acesso em: 07. fev. 2022.

PISEMSKAYA, N. B. El lenguaje y la teoría del caos. **Opcion [online]**, Maracaibo, v. 23, n. 53, p. 38-51, 2007. Disponível em: [http://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S1012-15872007000200004](http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1012-15872007000200004). Acesso em: 04 out.2021.

RODRIGUES, A. S. **A gramática do léxico: morfologia derivacional e o léxico mental**. München: LINCOM, 2015. Disponível em: <https://core.ac.uk/download/pdf/153413944.pdf> Acesso em: 04 out.2021.

SANTANA, N. M. O. **Estudo sócio-histórico-cognitivo das conceptualizações e categorizações do amor em cartas dos séculos XIX e XX**. 2019. Tese (Doutorado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2019.

SHANON, B. Fractal patterns in language. **New Ideas in Psychology**, v. 11, ed. 1, p. 105-109, mar., 1993. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/0732118X93900237>. Acesso em: 05 out. 2022.

SILVA T. D. L. da; SILVA, E. M. da. Mas o que é mesmo Corpus? – Alguns Apontamentos sobre a Construção de Corpo de Pesquisa nos Estudos em Administração. In: XXXVII ENCONTRO DA ANPAD, 37, 2013, Rio de Janeiro. **Anais [...]**. Maringá: Associação Nacional de Pós-Graduação e Pesquisa em Administração. 2013, p. 1-15. Disponível em: [http://www.anpad.org.br/admin/pdf/2013\\_EnANPAD\\_EPO1021.pdf](http://www.anpad.org.br/admin/pdf/2013_EnANPAD_EPO1021.pdf). Acesso em: 17 fev. 2022.

TEIXEIRA, J. O equilíbrio caótico do significado linguístico. **Diacrítica**, Braga, n.18, p. 189-207, 2004. Disponível em:  
<https://repositorium.sdum.uminho.pt/bitstream/1822/4335/1/caosSignific.pdf>  
Acesso em: 05 out. 2021.

ZAMORO AGUILAR, A. Teorías del caos y lingüística: aproximación caológica a la comunicación verbal humana. *Signa*. **Revista de la Asociación Española de Semiótica**, Biscaia, v. 21, p. 679-705, 2012. Disponível em:  
<https://revistas.uned.es/index.php/signa/article/view/6323/6056> Acesso em: 05 out. 2021.

DOMINGUES ALMEIDA, A. ARIADNE.  
NOTAS SOBRE O CORPUS EM ESTUDOS DA  
LINGUAGEM: A PERSPECTIVA FRACTAL E A  
TÉCNICA DA SATURAÇÃO TEÓRICA.  
**ENTREPALAVRAS**, FORTALEZA, v. 14, n. 1,  
E2737, p. 49-73, JAN.-ABR./20224 DOI:  
10.22168/2237-6321-12727