

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
CENTRO INTERDISCIPLINAR DE NOVAS TECNOLOGIAS NA EDUCAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA NA EDUCAÇÃO  
NÍVEL DOUTORADO

YGOR CORRÊA

Ação de Linguagem e Mineração de dados:  
níveis de complexidade na análise de textos

Porto Alegre  
2016

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
CENTRO INTERDISCIPLINAR DE NOVAS TECNOLOGIAS NA EDUCAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA NA EDUCAÇÃO  
NÍVEL DOUTORADO

YGOR CORRÊA

Ação de Linguagem e Mineração de dados:  
níveis de complexidade na análise de textos

Tese de Doutorado apresentada ao Programa de  
Pós-graduação em Informática na Educação da  
Universidade Federal do Rio Grande do Sul

Orientadora: Prof<sup>ª</sup>. Dra. Maria Cristina Villanova  
Biasuz

Coorientador: Prof. Dr. Eliseo Berni Reategui

Linha de Pesquisa: Interfaces Digitais em  
Educação, Arte, Linguagem e Cognição

Porto Alegre  
2016

CIP - Catalogação na Publicação

Corrêa , Ygor

Ação de Linguagem e Mineração de dados:níveis de complexidade na análise de textos / Ygor Corrêa . -- 2016.

151 f.

Orientadora: Maria Cristina Villanova Biasuz .  
Coorientadora: Eliseo Berni Reategui.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Centro de Estudos Interdisciplinares em Novas Tecnologias na Educação, Programa de Pós-Graduação em Informática na Educação, Porto Alegre, BR-RS, 2016.

1. Ação de linguagem. 2. Mineração de dados. 3. Sistemas Complexos Adaptativos. 4. Níveis de Complexidade. 5. Zonas de Desenvolvimento . I. Villanova Biasuz , Maria Cristina, orient. II. Berni Reategui, Eliseo, coorient. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).



SERVIÇO PÚBLICO FEDERAL  
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
CENTRO INTERDISCIPLINAR DE NOVAS TECNOLOGIAS NA EDUCAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA NA EDUCAÇÃO

**ATA SOBRE A DEFESA DE TESE DE DOUTORADO  
YGOR CORRÊA**

Às quatorze horas do dia vinte e quatro de março de dois mil e dezesseis, na sala 329 do PPGIE/CINTED, nesta Universidade, reuniu-se a Comissão de Avaliação, composta pelas Professoras Doutoradas: Margarete Axt, Dinorá Moraes de Fraga e Elisa Boff, para a análise da defesa de Tese intitulada “*Ação de linguagem e Mineração de dados: níveis de complexidade na análise de textos*”, do doutorando do Programa de Pós-Graduação em Informática na Educação Ygor Corrêa, sob a orientação da Profª. Drª. Maria Cristina Villanova Biasuz e coorientação do Prof. Dr. Eliseo Berni Reategui.

A Banca, reunida, após a apresentação e arguição, emite o parecer abaixo assinalado.

Considera a Tese aprovada

sem alterações;

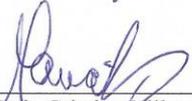
e recomenda que sejam efetuadas as reformulações e atendidas as sugestões contidas nos pareceres individuais dos membros da Banca;

Considera a Tese reprovada.

Considerações adicionais (a critério da Banca):

*A banca ressalta a excelência do trabalho, trabalhando com temas relacionados às mídias em suas interdisciplinas e operando com elas, potencializando as funcionalidades do recurso tecnológico usado.*

Porto Alegre, 24 de março de 2016.

  
\_\_\_\_\_  
Profª. Drª. Maria Cristina Villanova Biasuz  
Presidente e Orientadora

  
\_\_\_\_\_  
Prof. Dr. Eliseo Berni Reategui  
Coorientador

  
\_\_\_\_\_  
Profª. Drª. Margarete Axt  
PPGIE/UFRGS

  
\_\_\_\_\_  
Profª. Drª. Dinorá Moraes de Fraga  
UNIRITTER

*Videoconferência*   
\_\_\_\_\_  
Profª. Drª. Elisa Boff  
UCS

*“You might be poor,  
your shoes might be broken,  
but your mind is a palace.”  
Frank McCourt, 1996.*

## **AGRADECIMENTOS**

A conclusão desta tese não teria sido possível sem o apoio de algumas pessoas, bem como da instituição que foi fundamental durante todo o processo de doutoramento.

Agradeço, primeiramente, à minha família pelo apoio emocional, em especial, à minha mãe, Maria Corrêa, por ter me incentivado, desde muito pequeno, a estudar e acreditar que o mundo pode ser melhor por meio do conhecimento. Sem ela nada disso teria sido possível!

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) por ter me concedido a bolsa de doutorado.

À Universidade Federal do Rio Grande do Sul (UFRGS) por ter me propiciado momentos que se perpetuarão no tempo.

Ao Programa de Pós-Graduação em Informática na Educação (PPGIE) da Universidade Federal do Rio Grande do Sul (UFRGS), local em que pude amadurecer profissional e emocionalmente.

Agradeço à minha orientadora por ter estado comigo nesta empreitada, e, especialmente, ao Prof<sup>o</sup> Dr. Eliseo Berni Reategui, meu coorientador, pelo total apoio em minhas iniciativas acadêmicas.

Aos grupos de pesquisa N.E.S.T.A e GTech-Edu, assim como aos colegas e agora amigos Leonardo F. Baptista, Maristela Compagnoni Vieira, Isis da Costa, Daniel Epstein e, em especial, à Fúlvia da Silva Spohr, por ter sido extremamente importante ao longo de minha jornada neste curso, seja do ponto de vista emocional ou profissional. Muito obrigado a todo(a)s!

## RESUMO

Esta é uma pesquisa quali-quantitativa que tem por base teórico-epistemológica o Interacionismo Sociodiscursivo (ISD) e a Teoria da Complexidade, uma vez que ambas as teorias contribuem para a compreensão da ação de linguagem, enquanto Sistema Adaptativo Complexo (SAC), de sujeitos engajados em atividades de produção de textos. Dessa forma, na medida em que o desempenho de sujeitos de diferentes níveis de ensino tem sido apontado como problemático por órgãos de pesquisa em Educação, no que se refere às práticas de leitura e de produção textual, considerou-se relevante investigar a maneira como uma ferramenta digital de mineração de dados, neste caso, a ferramenta Sobek, pode evidenciar níveis de complexidade entre o conteúdo de um texto-referência e o texto-produzido. O objeto de estudo desta tese, portanto, é o recorte propiciado pela ferramenta Sobek, que ocorre por meio da geração de um grafo composto por palavras (signos) recorrentes nos textos, com vistas a sustentar a construção de uma perspectiva inter(trans)disciplinar, referente aos modos de extrair e interpretar o conteúdo desses a partir de níveis de complexidade. O *corpus* de pesquisa foi composto por 2 turmas de ensino superior, em nível de graduação, envolvendo a coleta de 34 textos. Esses textos foram submetidos à técnica de mineração de dados, com a utilização de uma nova versão da ferramenta, desenvolvida neste estudo, a qual permitiu importar e comparar, de forma automatizada, o conteúdo de dois textos, sendo possível gerar 3 tipos de grafos: de união, de intersecção e de diferença. O resultado da análise de dados indicou um padrão de estabilidade da ação de linguagem dos sujeitos, em relação ao conteúdo do texto-referência, sem apresentar significativa variabilidade de conceitos. Para além de uma nova versão da ferramenta, outra contribuição desta pesquisa foi a de sugerir que níveis de complexidade em caráter de estabilidade e variabilidade podem ser associados a zonas de desenvolvimento humano.

**Palavras-chave:** Ação de linguagem. Mineração de dados. Sistemas Complexos Adaptativos. Níveis de Complexidade. Zonas de Desenvolvimento

## ABSTRACT

This is a quali-quantitative research whose theoretical and epistemological basis are the Interactionism Sociodiscursivo (ISD) and the Complexity Theory, once both theories contribute to the understanding of the language action, as Complex Adaptive System (CAS), of individuals engaged in text production activities. Thus, as the individuals' performance in different levels of education has been identified as problematic by Research Institutes of Education, related to reading and textual production practices, it was considered relevant to investigate the way a digital data mining tool, in this case, Sobek, may reveal levels of complexity between the content of a reference-text and a produced-text. The object of study of this thesis, so is the cut provided by Sobek tool, which occurs through the generation of a graph composed of words (signs) recurring in the texts, in order to support the construction of an inter(trans) disciplinary perspective, referring to ways to extract and interpret the content of the texts. The research corpus was composed of 2 classes of higher education at the undergraduate level, involving the collection of 34 texts. These texts were submitted to data mining technique, with the use of a new version of the tool, developed in this study, which allowed to import and compare, automatically, the contents of two texts automatically, been possible to generate 3 types of graphs: union, intersection and difference. The result of data analysis indicated a recurring pattern in the analyzed texts, which was to maintain the language action as stable as possible to the content of the reference-text, without presenting significant variability of concepts. In addition to the new version of the tool, another contribution of this research was to suggest that the levels of complexity in stability and variability may be associated with human development zones.

**Keywords:** Language action. Data mining. Complex Adaptive Systems. Complexity levels. Development Zones

## LISTA DE FIGURAS

|   |     |
|---|-----|
| <b>Figura 1:</b> Interface da ferramenta Sobek .....                                      | 79  |
| <b>Figura 2:</b> Grafo ilustrativo da interface de geração e manipulação de dados .....   | 79  |
| <b>Figura 3:</b> Texto-referência minerado pelo professor.....                            | 81  |
| <b>Figura 4:</b> Texto-referência minerado pelo professor.....                            | 82  |
| <b>Figura 5:</b> Texto <i>Avaliação Terapêutica via Internet</i> .....                    | 87  |
| <b>Figura 6:</b> Texto <i>Terapia pela Internet</i> .....                                 | 87  |
| <b>Figura 7:</b> <i>Texto-discurso do Chat</i> .....                                      | 88  |
| <b>Figura 8:</b> Termo <i>paciente</i> adaptado no Chat .....                             | 91  |
| <b>Figura 9:</b> Termo <i>virtual</i> adaptado no Chat.....                               | 91  |
| <b>Figura 10:</b> Interface Sobek –Texto A (Texto Empírico).....                          | 94  |
| <b>Figura 11:</b> Interface Sobek – Botão de inserção do Texto B .....                    | 95  |
| <b>Figura 12:</b> Interface Sobek – Importação do TEXTO B .....                           | 95  |
| <b>Figura 13 -</b> Interface Sobek: Operações de grafos .....                             | 96  |
| <b>Figura 14 -</b> Interface Sobek: Tipos de Operações .....                              | 96  |
| <b>Figura 15:</b> Interface Sobek – Comparação com o texto.....                           | 97  |
| <b>Figura 16:</b> Interface Sobek – Layout Anterior .....                                 | 98  |
| <b>Figura 17:</b> Turma 1 - Grafo do texto-referência.....                                | 107 |
| <b>Figura 18:</b> Grafo conceito <i>Programa</i> e inter-relações .....                   | 108 |
| <b>Figura 19:</b> Turma 1 - Grafo conceito <i>Possibilidades</i> e subinter-relações..... | 109 |
| <b>Figura 20:</b> Nova versão da ferramenta Sobek .....                                   | 112 |
| <b>Figura 21:</b> Grafo geral T6.....   | 117 |
| <b>Figura 22:</b> Grafo geral T9.....   | 117 |
| <b>Figura 23:</b> Grafo geral T13.....  | 118 |
| <b>Figura 24:</b> Grafo de intersecção T6.....  | 119 |
| <b>Figura 25:</b> Grafo de intersecção T3.....  | 120 |
| <b>Figura 26:</b> Grafo de intersecção T10.....   | 121 |
| <b>Figura 27:</b> Grafo de diferença T13 .....  | 122 |
| <b>Figura 28:</b> Grafo de diferença T 2 .....  | 123 |
| <b>Figura 29:</b> Grafo de diferença T12 .....  | 124 |
| <b>Figura 30:</b> Turma 2 - Grafo do texto-referência.....                                | 127 |
| <b>Figura 31:</b> Turma 2 - Grafo conceito <i>Perspectivas</i> e inter-relações .....     | 128 |

|  |     |
|--|-----|
| <b>Figura 33:</b> Grafo geral T5.....            | 134 |
| <b>Figura 34:</b> Grafo geral T6.....            | 134 |
| <b>Figura 35:</b> Grafo geral T14.....           | 135 |
| <b>Figura 36:</b> Grafo de intersecção T6.....   | 136 |
| <b>Figura 37:</b> Grafo de intersecção T17 ..... | 137 |
| <b>Figura 38:</b> Grafo de intersecção T14.....  | 138 |
| <b>Figura 39:</b> Grafo de diferença T 4.....    | 139 |
| <b>Figura 40:</b> Grafo de diferença T17 .....   | 139 |
| <b>Figura 41:</b> Grafo de diferença T18.....    | 140 |

## LISTA DE TABELAS

|   |     |
|---|-----|
| <b>Tabela 1:</b> Termos minerados .....                                     | 89  |
| <b>Tabela 2:</b> Níveis de complexidade do Texto-referência (Turma 1) ..... | 110 |
| <b>Tabela 3:</b> Níveis de complexidade do Texto-referência (Turma 2) ..... | 129 |

## LISTA DE GRÁFICOS

|   |     |
|---|-----|
| <b>Gráfico 1:</b> Níveis 1, 2 e 3 de complexidade nos textos-empíricos (Turma 1)..... | 114 |
| <b>Gráfico 2:</b> Níveis de Intersecção e Diferença no texto-empírico (Turma 1).....  | 115 |
| <b>Gráfico 3:</b> Comparação Nível 1 – texto-referência e textos-empíricos .....      | 125 |
| <b>Gráfico 4:</b> Comparação Nível 2 – texto-referência e textos-empíricos .....      | 125 |
| <b>Gráfico 5:</b> Comparação Nível 3 – texto-referência e textos-empíricos .....      | 126 |
| <b>Gráfico 6:</b> Níveis 1, 2 e 3 de complexidade nos textos-empíricos (Turma 2)..... | 131 |
| <b>Gráfico 7:</b> Níveis de Intersecção e Diferença no texto-empírico (Turma 2).....  | 132 |
| <b>Gráfico 8:</b> Comparação Nível 1 – texto-referência e textos-empíricos .....      | 141 |
| <b>Gráfico 9:</b> Comparação Nível 2 – texto-referência e textos-empíricos .....      | 142 |
| <b>Gráfico 10:</b> Comparação Nível 3 – texto-referência e textos-empíricos .....     | 142 |

# SUMÁRIO

|          |  |            |
|----------|--|------------|
| <b>1</b> | <b>INTRODUÇÃO .....</b>  | <b>14</b>  |
| 1.1      | Questão de pesquisa.....   | 19         |
| 1.2      | Objetivos de pesquisa .....  | 20         |
| 1.2.1    | Objetivo Geral .....   | 20         |
| 1.2.2    | Objetivos específicos .....  | 21         |
| 1.3      | Organização da tese .....  | 21         |
| <b>2</b> | <b>PRESSUPOSTOS TEÓRICOS.....</b>                                  | <b>25</b>  |
| 2.1      | A busca por uma abordagem inter(trans)disciplinar.....             | 26         |
| 2.2      | Interacionismo Sociodiscursivo (ISD).....                          | 30         |
| 2.2.1    | A concepção teórica do ISD .....                                   | 31         |
| 2.2.2    | Signo linguístico, atividade de linguagem e ação de linguagem..... | 37         |
| 2.2.3    | Da produção de textos-empíricos à mediação formativa.....          | 44         |
| 2.3      | A Teoria da Complexidade na Linguística Aplicada.....              | 47         |
| 2.3.1    | Complexo, dinâmico, adaptativo e não linear .....                  | 52         |
| 2.3.2    | Caótico, imprevisível, sensível às condições iniciais .....        | 54         |
| 2.3.3    | Aberto, sensível ao <i>feedback</i> .....                          | 55         |
| 2.3.4    | Atratores estranhos e fractais.....                                | 56         |
| 2.3.5    | Auto-organização/Emergência .....                                  | 57         |
| 2.4      | Os princípios de Vygotsky à luz da Teoria da Complexidade.....     | 62         |
| 2.5      | A ferramenta de mineração de dados Sobek.....                      | 76         |
| <b>3</b> | <b>METODOLOGIA .....</b>   | <b>83</b>  |
| 3.1      | Estudo Preliminar .....  | 85         |
| 3.2      | O desenvolvimento de funções .....                                 | 93         |
| 3.3      | Estudo Final .....   | 98         |
| <b>4</b> | <b>CONSIDERAÇÕES FINAIS .....</b>                                  | <b>143</b> |
|          | <b>REFERÊNCIAS .....</b>   | <b>147</b> |

# 1 INTRODUÇÃO

A presente pesquisa está pautada no desenvolvimento humano, por meio da atividade de linguagem (escrita ou oral), entendendo-a sob a forma de processos sociodiscursivos que se materializam em textos escritos, mediante ação de linguagem dos sujeitos (BRONCKART, 2006, 2008). Dessa maneira, ao levar em consideração que boa parte dos sujeitos, em todos os níveis de ensino, tem dificuldades ao ler, compreender e escrever textos, dada a complexidade da dinâmica dialógica e polifônica da língua (BAKHTIN, 2003, 2006), percebeu-se tal aspecto como uma problemática a ser investigada, que comporta, em si, implicações de relevância social, política e educacional. Relativo a isso, Marcuschi (2008, p.230) traz como prova o fato de que “a não naturalidade da boa compreensão é o resultado da avaliação internacional feita pelo PISA em 41 países pela Organização para Cooperação e Desenvolvimento Econômico (OECD-PISA)”. Referindo-se à aplicação desse exame, Marcuschi (2008, p.230), embora critique a qualidade de exames dessa ordem, aponta para o fato alarmante de que “o Brasil ficou entre os últimos cinco países com nota inferior a 4 pontos numa escala de 1 a 8”. O que se infere, a partir dos resultados deste teste, assim como pelo teste do Sistema Nacional de Avaliação da Educação Básica (SAEB), é que a compreensão do que é lido gira em torno apenas de 50-60% da expectativa de domínio de leitura e produção textual.

Com o objetivo de afirmar a possível origem de tal dificuldade, Marcuschi (2008, p.231) considera que “isto se deve à própria natureza da linguagem, que não é transparente nem funciona como uma fotografia ou xerox da realidade”. Embora, os apontamentos de Marcuschi (2008), quanto aos índices, datem de um período anterior ao do presente estudo, os mesmos órgãos de pesquisa em desenvolvimento e educação continuam a apontar baixos níveis de desempenho para as atividades que abrangem ações situadas na área das Linguagens, mais especificamente, Língua Portuguesa. Desse modo, ao considerar a constituição interacional da língua em que “os sujeitos são vistos como atores/construtores sociais, sujeitos ativos que – dialogicamente – se constroem e são construídos no texto” (KOCH e ELIAS, 2012, p.10), torna-se relevante a busca por uma interpretação científica da utilização de ferramentas digitais que possam ser percebidas como mediadoras da produção de conhecimentos (LÉVY, 1999), as quais,

talvez, possam auxiliar os sujeitos em uma melhor interpretação da complexidade inerente a suas práticas de produção de textos.

Nesse sentido, a premissa inicial deste estudo é a de ampliar o diálogo inter(trans)disciplinar entre Informática, Educação e Teoria da Complexidade, quanto à utilização de tecnologias digitais (LÉVY, 1999), como mediadoras de processos de ensino e aprendizagem em contexto educacional. Esta tomada de posição está em consonância com o próprio objetivo do Programa de Pós-Graduação em Informática na Educação (PPGIE), da Universidade Federal do Rio Grande do Sul (UFRGS), que é o de propor uma tensão coerente entre duas ciências ou mais, a fim de gerar subsídios científicos, que permitam refletir, construir e aplicar construtos teóricos à produção de conhecimento *para e a partir de* recursos tecnológicos que visem ao desenvolvimento humano do ponto de vista físico, cognitivo e emocional (AXT *et. al*, 2011).

Dessa maneira, diante da problemática desencadeadora desta pesquisa e percebendo a existência de recursos digitais para o apoio à leitura e à produção de textos (MACEDO *et. al*, 2011), como é o caso da ferramenta de mineração de dados Sobek, surge a demanda científica a ser explorada pelo autor desta tese que, enquanto linguista, se propõe a estabelecer uma linha de convergência em conjunção entre a Teoria da Complexidade e o Interacionismo Sociodiscursivo (ISD), no campo da Informática na Educação. Esta abordagem, por sua vez, abarca o ISD no sentido de compreender os modos de estruturação e as condições de funcionamento da linguagem, entendendo que ela é o instrumento fundamental da gnosologia e da praxiologia humana (BRONCKART, 2007). Uma vez que para o ISD a atividade humana é mediada por instrumentos (signos) e as ações dos sujeitos orientam-se por esses, que ao serem significados materializam-se, então a língua é um instrumento, forjado, na interação de grupos sociais. Tendo isso em vista, cabe discutir e teorizar acerca dos modos de interpretar a complexidade do conteúdo de textos empíricos submetidos ao processo de mineração de dados, denominando a ação de linguagem como um Sistema Adaptativo Complexo (SAC), a partir das postulações da Teoria da Complexidade, situada na Linguística Aplicada. Partindo desta premissa, considera-se que a Teoria da Complexidade pode alicerçar perspectivas outras, que não aquelas já exploradas de modo isolado dentro do discurso de cada ciência (AXT *et. al*, 2011; MORIN, 2008; LARSEN-FREEMAN e CAMERON, 2008). O estudo, contudo, não visa a explorar o conteúdo de textos empíricos, em caráter puramente linguístico e disciplinar, mas se propõe a desenvolver uma compreensão do caráter fenomenológico e dinâmico da ação

de linguagem dos sujeitos de pesquisa na tentativa de compreender se o conteúdo apontado por um processo de mineração de dados, via ferramenta digital, pode, de fato, evidenciar níveis de complexidade próximos ou afastados de zonas de desenvolvimento (VYGOSTKY, 1998) em relação à atividade de produção textual.

Muito embora o objetivo deste estudo não seja o de avaliar o desenvolvimento linguístico dos sujeitos (ação de linguagem), de modo a conceituar noções de gêneros discursivos – primário ou secundário (BAKHTIN, 2003, 2006) –, tipos de discurso – narrativo, argumentativo, descritivo, etc. (BRONCKART, 2006, 2008) – ou gêneros textuais – anúncio de jornal, receita de bolo, artigo científico (MARCUSCHI, 2008), cabe ressaltar que, sem dúvida, a apropriação dessas noções está subentendida nesta pesquisa, no que diz respeito aos modos de produzir textos. Referente a isso, afirma-se que os textos empíricos, submetidos ao processo de mineração, neste estudo, originam-se de gêneros discursivos secundários, sendo todos da esfera acadêmica (artigos científicos e capítulos teóricos de livros) e produzidos em contexto educacional de ensino superior, que “aparecem em circunstâncias de uma comunicação cultural mais complexa e relativamente mais evoluída, principalmente escrita, artista, científica, sociopolítica” (BAKHTIN, 2003, p.281). Conquanto práticas de linguagem dizem respeito “às dimensões particulares do funcionamento da linguagem em relação às práticas sociais, em geral, tendo a linguagem função de mediação em relação a essas últimas” (SCHNEUWLY e DOLZ, 2004, p.72), é por meio dessas práticas que a atividade de linguagem “transforma-se em forma de expressão do pensamento, da experiência ou da percepção” (SCHNEUWLY e DOLZ, 2004, p.76), materializada na forma de textos empíricos.

Então, o que talvez justifique a dificuldade de transição do gênero primário que “se constitui em circunstâncias de comunicação verbal espontânea” para o gênero secundário – mais complexo e, por vezes, destituído de motivações pessoais para sua construção – seja a transição de uma composição para outra, a qual tende a evidenciar a complexidade do processo e, além disso, pode ser encarada como uma tensão entre sistemas distintos, gerando, assim, condições de desenvolvimento para os sujeitos (SCHNEUWLY e DOLZ, 2004). Neste tocante, vê-se a construção de textos escritos, concretos e complexos da esfera acadêmica como uma maneira de observar operações cognitivas de linguagem na forma de manifestação de conceitos (signos linguísticos) para evidenciar relações de implicação (estabilidade) e autonomia (instabilidade) dos sujeitos ao conteúdo que produzem. Concebe-se, portanto, o papel central da linguagem

no desenvolvimento humano, quanto à interiorização de formas superiores do pensamento, ou seja, da construção de conhecimento (VYGOSTKY, 1998). Cabe ainda ressaltar que, por se considerar que “o sentido não está no leitor, nem no texto, nem no autor, mas se dá como efeito das relações entre eles e das atividades desenvolvidas” (MARCUSCHI, 2008) é que se adota a Teoria da Complexidade como aporte teórico capaz de sustentar o desafio de observar a tessitura complexa do que emerge de práticas de linguagem, que têm textos empíricos como instrumento de concretização da ação de linguagem no nível coletivo e individual.

Ao assumir a atividade de linguagem, mais precisamente a ação de linguagem, sob a constituição de textos escritos, infere-se, segundo Marcuschi (2008), que os textos se apresentam como “sistemas instáveis e sua instabilidade é sempre um estado transitório de adaptação a um determinado objetivo e contexto (p. 243)”. Sob essa ótica, compreende-se que produções discursivas não são produtos fixos, mas um agir linguageiro quase sempre emergente, oriundo das condições iniciais oferecidas (textos externos à ação de quem os interpreta) aos sujeitos em diálogo com suas manifestações linguísticas, em uma relação dialógica e polifônica (BAKHTIN, 2003, 2006). Posto isso, tanto língua, quanto linguagem devem ser compreendidas como “atividade social, histórica e cognitiva, desenvolvida por práticas socioculturais” (MARCUSCHI, 2008, p. 64). Essa afirmação é basilar para o diálogo que está estabelecido neste estudo, de que a linguagem não possui regras fixas e imutáveis, mas que é, por excelência, socialmente construída, tendo em vista a situação em que se encontram os sujeitos pertencentes à determinada comunidade de fala.

No que concerne ao uso de recursos digitais de mineração e à possível identificação de níveis de complexidade, não da transição de um gênero primário para um gênero secundário, mas sim, das relações transicionais entre dois gêneros secundários, como é o caso deste estudo, quando da leitura de um texto-referência<sup>1</sup>, indicado por um professor, em contexto de ensino superior, para que o aluno produza seu texto (texto-empírico), adotou-se a ferramenta Sobek (REATEGUI, PINHO e EPSTEIN, 2011) no intuito de investigar as relações entre o conteúdo dos textos envolvidos (texto-referência e texto-empírico) e refletir a respeito do tipo mediação proporcionada. Os textos abarcados por esta pesquisa dividem-se em dois grupos: texto-

---

<sup>1</sup> Neste estudo, entende-se que tanto o texto-referência, quanto o texto-empírico são textos empiricamente constituídos pela ação de linguagem de seus respectivos autores, porém para fins de operacionalização das proposições apresentadas nesta pesquisa, optou-se por denominá-los como referido.

referência, ou seja, um texto que não foi produzido pelo atuante (sujeito), denominando pelo ISD de texto-externo; e texto-empírico, que diz respeito à instância participativa dos agentes (sujeitos).

Relativo à ferramenta de mineração Sobek, esta foi desenvolvida por um grupo de pesquisa multidisciplinar da UFRGS, sob a coordenação do Prof<sup>o</sup> Dr. Eliseo Berni Reategui e tem desencadeado pesquisas em nível de graduação, especialização, mestrado e doutorado, a partir de diálogos estabelecidos com áreas do conhecimento, tais como Educação, Psicologia e Linguística. Sobek, que é descrita no seu respectivo capítulo, por sua vez, está implicada tanto na abordagem teórica, quanto na metodológica deste estudo, por duas razões: a) na teórica – por pautar a proposta investigativa do estudo a partir do modo operacional como ocorre o processo de mineração, pois nasce do recorte feito pela própria técnica (de mineração), como condição inicial para compreender a complexidade processual de textos-empíricos e b) na metodológica – por utilizá-la como método de extração de dados, o qual será contextualizado ao aporte teórico adotado.

Estudos realizados com a ferramenta Sobek (REATEGUI, PINHO e EPSTEIN, 201; MACEDO *et. al*, 2011) remetem a benefícios relacionados aos processos de leitura e escrita: a) minerar um texto e obter, na forma de grafos (conceitos relacionados ou não entre si), um panorama de conceitos que são recorrentes e que, na proposta de mineração, revelam especificidades do conteúdo do texto lido para que tarefas outras sejam pautadas na inferência de que é relevante levar em consideração o que está exposto no grafo; b) minerar um texto, seguindo a mesma lógica descrita no item “a”, porém com a finalidade de, a partir do que fora exposto, conceber os conceitos para uma produção escrita centrada no texto que serviu como referência para a produção textual. Por outro lado, a proposição deste estudo visa a progredir nos modos de realizar o processo e a compreensão da ação de linguagem e de mineração de dados, pois se destina à observação da dinâmica composicional entre dois textos, e não, de um processo minerativo isolado, onde apenas um texto está em questão. Além disso, a proposta de pesquisa está focada no estudo da dinamicidade do conteúdo, e não, no seu estado de fixidez, a fim de que, além do conteúdo, possa-se postular sua constituição relativa a zonas de desenvolvimento. A seguir, apresenta-se a questão de pesquisa acompanhada de possíveis hipóteses, assim como os seus objetivos.

## 1.1 Questão de pesquisa

Em relação à base teórico-epistemológica, apresentada brevemente na introdução desta pesquisa, acredita-se que investigar a utilização de uma ferramenta de mineração de dados, a qual fornece um panorama das relações recorrentes em textos-empíricos, na forma de grafo de conceitos, talvez, possa auxiliar na identificação e compreensão da complexidade da ação de linguagem dos sujeitos. Assim, cogita-se que a observância de inter-relações entre o texto-referência (externo à ação dos sujeitos/atuantes) e o texto-empírico (produzido em uma atividade específica e dotado de motivações e intenções pessoais) possa relevar níveis de complexidade ou, ainda, diferentes estratégias para minerar textos. Essa observação, talvez, possa gerar problematizações e contribuições científicas de cunho tecnológico, educacional e linguístico para um melhor entendimento do desenvolvimento humano, pautado na atividade de linguagem e suas implicações com alguns conceitos, como o de Zona de Desenvolvimento Proximal (ZDP) (VYGOSTKY, 1998). Acredita-se, portanto, que textos-empíricos submetidos à mineração de dados possam propiciar um panorama das características emergentes do processo de construção de conhecimento, assim como dos modos de realizar o próprio processo de extração. Dessa forma, a questão principal de pesquisa é esta:

*De que maneira uma ferramenta de mineração de dados pode auxiliar na análise do conteúdo de uma produção textual, de modo a revelar níveis de estabilidade/variabilidade desta em relação ao conteúdo de um texto-referência?*

Diante desse questionamento, o pressuposto de pesquisa é o de que uma ferramenta de mineração pode contribuir para evidenciar a complexidade linguística inerente à ação de linguagem, de maneira a conduzir a ação investigativa a características que, na perspectiva da Teoria da Complexidade, são mais do que conceitos apresentados de forma estável e isolada, os quais representam a produção de conhecimento dos sujeitos engajados em produções textuais. Dessa forma, se níveis de complexidade forem evidenciados, assim como a maneira de observá-los, a partir do aporte teórico adotado neste estudo, então infere-se que, deste modo, será possível contribuir para com um modelo apreciativo de mineração de textos-empíricos, que

potencializa a identificação do conteúdo se comparado a uma observação “a olho nu”, sem auxílio de uma ferramenta digital.

No entanto, partindo dos pressupostos mencionados, será necessário aprofundar os aportes teóricos adotados, quais sejam: a) o desenvolvimento humano, nesta tese estudado enquanto práticas de produção textual, à luz da Teoria da Complexidade; b) a prática de mineração de dados e a identificação do recorte feito por esse processo em relação ao conteúdo de produções textuais derivadas de contexto de ensino e aprendizagem; c) a concepção de níveis de complexidade que podem evidenciar aspectos e modos de minerar, e a possibilidade de gerar grafos distintos. Esses podem permitir ao usuário observar implicações associadas à atividade de produção de textos, bem como associar os resultados às zonas de desenvolvimento de quem os produz.

Diante do pressuposto elaborado, são apresentadas considerações que podem nortear instâncias do conhecimento, enquanto estratégias investigativas, tais como estas:

a) Como a inserção de técnicas de mineração de dados no processo de identificação do conteúdo de atividades de linguagem, produções textuais, pode contribuir para a compreensão de sub e inter-relações presentes em textos-empíricos?

b) Quais são as demandas deste processo, diante da programação de uma ferramenta frente à complexidade linguística manifestada *na* e *pela* linguagem utilizada pelos sujeitos?

c) Que elementos podem evidenciar a emergência de novos padrões e/ou comportamentos, os quais podem caracterizar textos-empíricos como Sistemas Adaptativos Complexos (SACs)?

Considera-se que, frente às hipóteses e postulações presentes nesta pesquisa, surjam novas frentes investigativas como resultado deste fazer científico inter(trans)disciplinar.

## **1.2 Objetivos de pesquisa**

A seguir são descritos os objetivos investigados nesta pesquisa.

### **1.2.1 Objetivo Geral**

Investigar a maneira como uma ferramenta de mineração de dados pode auxiliar na análise do conteúdo de uma produção textual, de modo a revelar seus níveis de estabilidade ou variabilidade em relação ao conteúdo de um texto referência.

### 1.2.2 Objetivos específicos

- Analisar a utilização da ferramenta de mineração de dados Sobek no que diz respeito ao nível de complexidade propiciado pelo recorte por ela oferecido, na forma de grafos, ao submeter textos-empíricos ao processo de extração de dados.
- Depreender possíveis níveis de estabilidade e variabilidade do conteúdo do texto-empírico, considerando sua dinamicidade composicional em relação ao texto-referência de uma atividade de produção textual, tendo por base conceitos presentes na Teoria da Complexidade, ao considerar textos-empíricos como Sistemas Adaptativos Complexos (SACs).
- Elaborar, a partir das (im)possibilidades da ferramenta, formas de pensar a geração de tipos de grafos automatizados e distintos acerca de inter-relações entre dois textos, sendo esses capazes de apontar para diferentes níveis de complexidade do conteúdo exposto à mineração.
- Discutir o processo de mineração de dados, os níveis de complexidade identificados nos textos-empíricos e suas características, a fim de evidenciar o quanto eles sugerem um padrão de desenvolvimento humano, nesta tese, baseado no conceito de ação de linguagem.

Uma vez pontuados os objetivos, a seguir, é descrita a organização dos capítulos que constituem esta pesquisa.

## 1.3 Organização da tese

Este estudo situa-se na linha de pesquisa de Interfaces Digitais em Educação, Arte, Linguagem e Cognição e está vinculado ao grupo de pesquisa intitulado Núcleo de Estudos em Subjetivação, Tecnologia e Arte (N.E.S.T.A), do Programa de Pós-Graduação em Informática na Educação (PPGIE), da Universidade Federal do Rio Grande do Sul (UFRGS), sob a coordenação da Prof.<sup>a</sup> Dr.<sup>a</sup> Maria Cristina Villanova

Biasuz, assim como ao Grupo de Pesquisa em Tecnologia Aplicada à Educação (GTech-Edu), coordenado pelo Prof. Dr. Eliseo Berni Reategui. Esta pesquisa está dividida em duas dimensões, a saber: a primeira, acerca das postulações teórico-epistemológicas, norteadoras do estudo realizado e, em seguida, a segunda, caracterizada pela dimensão prática e aplicada de análise e discussão do objeto de estudo investigado, neste caso, textos-empíricos submetidos ao processo de mineração de dados. A tese é composta por 4 capítulos, que constituem a totalidade do estudo.

No *capítulo 1, Introdução*, apresentou-se a problemática desencadeadora do tema desta pesquisa em diálogo com a justificativa para sua realização, seguido da perspectiva de rompimento ao Paradigma Simplicista, de modo a construir produção de conhecimento em caráter inter(trans)disciplinar.

O *capítulo 2, Pressupostos teóricos*, por sua vez, é constituído, inicialmente, pela apresentação da delimitação do objeto de estudo da pesquisa que é a compreensão das implicações derivadas da utilização de uma ferramenta de mineração de dados para a interpretação da complexidade processual presente na relação do conteúdo de dois textos. Dessa maneira, busca-se pontuar, em linhas gerais, o que se está propondo como convergência das teorias que são exploradas na progressão do desenvolvimento do estudo. Ainda neste capítulo, discorre-se, em *2.1*, sobre *A busca por uma abordagem inter(trans)disciplinar*, situada no fazer científico da operacionalização do pensamento complexo de cunho filosófico-educacional (NICOLESCU, 1999; MORIN, 2008), quando de um diálogo unificado a partir de ciências distintas, subentendendo a adesão desse por ciências de núcleo *hard* (Informática, Física, etc.) em relação às ciências humanas (Educação, Linguística, etc.).

Na continuidade da proposição de pensar o desenvolvimento humano por meio da atividade/ação de linguagem, ou seja, por interações sociodiscursivas, apresenta-se o subcapítulo *2.2, Interacionismo Sociodiscursivo (ISD)*, como teoria oriunda das Ciências Sociais Humanas (Psicologia, Educação e Linguística), capaz de situar o que se compreende por processos desenvolvimentais pautados em atividades centradas na produção de linguagem, via textos escritos e empíricos (BRONCKART, 2006, 2008). Na seção *2.2.1*, discorre-se sobre *A concepção teórica do ISD*, seguido de *2.2.2 Signo linguístico, atividade de linguagem e ação de linguagem* e *2.2.3 Da produção de textos-empíricos à mediação formativa*. A aderência aos postulatos do ISD deve-se ao fato de que tal linha teórica pode auxiliar no entendimento de aspectos relacionados à produção textual, à noção de signo linguístico e ao desenvolvimento humano como um todo.

No subcapítulo 2.3, *Teoria da Complexidade na Linguística Aplicada*, descreve-se o conceito de Complexidade, seguido da caracterização do que são Sistemas Adaptativos Complexos (SACs), na medida em que este estudo considera o ser humano e a atividade de linguagem, na forma de textos escritos empíricos, enquanto sistemas vivos, dinâmicos e em constante adaptação, sendo suscetíveis à manifestação de padrões complexos, quando expostos à determinadas atividades. Esse posicionamento teórico visa a contribuir para a interpretação do recorte feito pela ferramenta de mineração, que é o de apresentar um panomarama do todo textual, na forma de grafo dotado de conceitos recorrentes no texto. A aplicabilidade da Teoria da Complexidade, portanto, gira em torno da possibilidade de discorrer acerca de possíveis padrões que possam ser identificados no texto-produzido em relação ao conteúdo do texto-referência, tratamento metodológico este que visa a problematizar tendências desenvolvimentais presentes na atividade de linguagem dos sujeitos desta pesquisa.

Na sequência, o subcapítulo 2.4, *Os princípios de Vygostky à luz da Teoria da Complexidade*, destina-se à construção de uma aproximação dos pressupostos da Teoria da Complexidade aos construtos vygostkianos, a saber, como forma de evidenciar que suas contribuições, embora não fossem pautadas na complexidade, a ela remetem de forma bastante notável. Em verdade, esse subcapítulo, aborda tal ótica a fim de, após realizar a análise de dados minerados, conjecturar sobre o fato de possíveis tendências desenvolvimentais, identificadas por meio de textos-empíricos, poderem parecer mais estáveis ou variáveis em relação à Zona de Desenvolvimento Real (ZDR) e à Zona de Desenvolvimento Proximal (ZDP). Não se tratará da identificação dessas zonas, mas dos desdobramentos e possíveis modos de operacionalizar a produção de conhecimento a partir dos resultados obtidos em relação a elas. Além disso, a aproximação teórica proposta visa a progredir na instauração do pensamento complexo, no que tange aos conceitos de desenvolvimento e aprendizagem.

O subcapítulo 2.5, *Ferramenta de Mineração de Dados Sobek*, descreve as funcionalidades desse recurso digital e estudos realizados a partir de sua aplicação, em contextos empíricos de ensino e aprendizagem, quanto as ações de leitura e produção de textos. Além disso, quando da referência a estudos anteriores, é explicitado o modo como essa tecnologia pode servir de instrumento de potencialização e mediação de práticas discursivas referentes a atividades que envolvem a ação de linguagem dos sujeitos e reverberam na concretizam do pensamento consciente como forma de desenvolvimento humano.

O *capítulo 3, Metodologia*, apresenta a pesquisa conduzida, sendo ela uma pesquisa quali-quantitativa que atenta para as ações de linguagem, concretizadas via textos empíricos, assim como para índices, identificados pela técnica de mineração como forma de evidenciar se o conteúdo do texto-empírico mantém-se estável (dados iguais ao texto-referência) ou variável (dados diferentes). O cenário de coleta de dados é abordado como forma de situar a procedência dos textos-empíricos coletados, o número de turmas de nível de ensino superior e as estratégias metodológicas concretizadas no estudo. Ainda neste capítulo, aborda-se, inicialmente, a realização do 3.1 *Estudo Preliminar* realizado com a ferramenta Sobek e os seus desdobramentos, quando da tentativa de identificação da complexidade do conteúdo do texto-empírico em relação ao do texto-referência.. A partir desse estudo e das implicações constatadas, propõe-se 3.2. *O desenvolvimento de novas funções*, de forma a minerar dois textos simultaneamente, a fim de apresentar três tipos distintos de grafos: de união, de intersecção e de diferença. Após essa etapa, os textos coletados são submetidos a nova versão, que suporta e compara dois textos simultaneamente, de maneira a permitir a identificação da complexidade inerente às relações presentes em cada texto.

No subcapítulo 3.3 *Estudo Final*, são apresentados os desdobramentos referentes aos conteúdos dos textos-empíricos submetidos à técnica de mineração de dados da nova versão da ferramenta e, com isso, constata-se a predominância de tendências na ação de linguagem dos sujeitos, prevalecendo a ação de manter-se estável ao conteúdo do texto-referência (externo), ou seja, fazer uso da maior parte dos conceitos que serviram de condição inicial para a produção do texto-empírico. Essa postulação é corroborada pelos índices apresentados e serve como reverberação para problematizar o modo como os resultados minerados, a partir de tipos distintos de grafos, podem ser pensados em relação ao conceito de mediação formativa (ISD), bem como às zonas de desenvolvimento vygostkianas.

Por fim, no *capítulo 4, Considerações finais*, pontuam-se as contribuições deste estudo para o campo da Informática na Educação e novas frentes investigativas percebidas ao longo da realização do presente estudo. A seguir, no próximo capítulo, são apresentados os pressupostos teóricos que norteiam esta pesquisa.

## 2 PRESSUPOSTOS TEÓRICOS

Neste capítulo, apresenta-se a base teórico-epistemológica deste estudo, primeiro propondo uma abordagem inter(trans)disciplinar, justamente por compreender que os resultados gerados por meio de uma ferramenta de mineração de dados, como é o caso de Sobek, não devem ser concebidos única e exclusivamente pela perspectiva da Informática, de modo a não induzir sua utilização apenas a uma ótica disciplinar, mas conjecturar um diálogo com outras ciências e teorias. Nesse sentido, ao elencar a Teoria da Complexidade, situada na Linguística Aplicada, para analisar mais a fundo a tessitura complexa da consistência da ação de linguagem, denominando-a como um Sistema Adaptativo Complexo (SAC), além dela o humano e a língua/linguagem também estão compreendidos como SACs (MORIN 1996, 2000, 2008, 2009; BRONCKART, 2006, 2008; LARSEN-FREEMAN e CAMERON, 2008).

Assim, a partir dessa premissa, infere-se que a técnica de mineração de dados pode tornar observável uma parte do todo composicional do texto, por meio de grafos, permitindo que seu usuário obtenha uma ótica fractalizada em caráter estático, ou seja, um recorte pautado no isolamento de palavras/conceitos recorrentes, o qual pode apontar para um determinado estado do texto, enquanto SAC. Nesse sentido, a noção de recorte pode ser considerada como um panorama/paisagem (*landscape*) do todo textual, como propõem Larsen-Freeman e Cameron (2008), ao afirmarem que qualquer SAC, em estabilidade ou variabilidade interacional, apresenta estados aparentemente fixos que, na verdade, são transitórios, dada a dinamicidade presente em sua composição interna, assim como em relação à influência de fatores externos. Diante disso, é possível considerar que uma ferramenta de mineração de dados tem por função “capturar” conceitos que, embora, a princípio, pareçam dotados da mesma potência de expressão semântica, na verdade, são arbitrários na relação significado-significante, pois a produção de sentido de um conceito só pode ser observada sob sua condição enunciativa, socialmente construída e individualmente aplicada e validada à determinada prática discursiva.

Relativo a isso, ao propor a observação do conteúdo de uma produção textual, na relação com o texto-referência, considera-se relevante, com o auxílio de Sobek, analisar práticas languageiras por meio da ação de linguagem dos sujeitos, de maneira a identificar o que, na constituição do texto-empírico, se assemelha ou se distancia do que

é recorrente no texto-referência. Embora a condição linguística dos conceitos esteja sendo entendida como abstrata, uma vez que se está observando-os isoladamente, sem atribuir-lhes sua potência enunciativa onde foram aplicados, considera-se que a identificação do seu índice de recorrência tende a sugerir, do ponto de vista discursivo, o quanto os sujeitos se apropriam ou não de conceitos presentes no texto lido, replicando-os, como forma de manter a constituição do que lhes serviu como referência (condição inicial para a produção do texto-empírico).

Nesse tocante, a Teoria da Complexidade pode auxiliar na instauração de um *modus operandi*, no que se refere à compreensão da complexidade do conteúdo que se pode extrair de textos-empíricos, quando da utilização da ferramentada de mineração de dados Sobek, no intuito de propor a identificação de possíveis níveis de complexidade presentes em produções textuais. Relativo a isso, entende-se que o principal desafio está centrado em analisar o resultado do processo de extração de dados, verificando o que esse revela para quem o utiliza, e não se preocupando em categorizar, do ponto de vista linguístico, o conteúdo dos textos coletados, mas procurando possibilitar diferentes perspectivas de mineração. Com isso, pode-se apontar para tendências de desenvolvimento que podem gerar estratégias, para que o sujeito possa ir além do conteúdo do texto-referência (texto-externo), sendo assim capaz de progredir em sua construção de conhecimento.

Dessa forma, este estudo não comporta modelos prescritivos que demandem o enquadramento ou a descomplexificação dos processos humanos e linguísticos em seus formatos; pelo contrário, atenta para a interação entre dois sistemas - o texto-referência e o texto-empírico -, visando à explicação de fluxos desenvolvimentais e emergências linguísticas. Este capítulo está dividido cinco subcapítulos: 2.1 A busca por uma abordagem inter(trans)disciplinar; 2.2 Interacionismo Sociodiscursivo (ISD); 2.3 Teoria da Complexidade na Linguística Aplicada; 2.4 Os princípios de Vygostky à luz da Teoria da Complexidade e 2.5 A ferramenta de mineração de dados Sobek.

## **2.1 A busca por uma abordagem inter(trans)disciplinar**

Este subcapítulo discorre a respeito da busca por uma abordagem inter(trans)disciplinar no fazer científico, de modo a construir uma linha de raciocínio

em que ciências de núcleos distintos possam ser percebidas como convergentes e não antagônicas, no que se refere à observação de objetos de estudo de natureza complexa, sem que eles tenham que ser descomplexificados para que possam ser estudados apenas em caráter disciplinar. Para pôr em pauta essa questão, no que diz respeito à própria abordagem deste estudo, a seguir, são apresentadas algumas contribuições de Morin (2008) e Nicolescu (1999) acerca do Paradigma da Simplificação, na tentativa de instaurar uma das premissas presentes no pensar o complexo, sugerido por Morin (2008).

No horizonte proposto e apoiado nas postulações de Morin (2008) e Nicolescu (1999), as abordagens cartesianas são entendidas como aquelas que aparam as arestas de seus objetos de estudo, por meio do isolamento de um parâmetro ou mais e buscam descrições passíveis de serem sempre defendidas a partir de modelos pré-estabelecidos, os quais estão calcados em métodos de replicabilidade, sob vários ângulos, de forma a atingir sempre o mesmo resultado. Na perspectiva adotada por este estudo, essas são abordagens, inquestionavelmente, reducionistas diante de processos desenvolvimentais humanos e da atividade e ação de linguagem que exercem em práticas sociais, posto que não se dedicam a uma gama complexa de fatores relacionados a processos, mas a estados supostamente fixos. Dessa maneira, de acordo com Morin (1996), negar a complexidade inerente à natureza do humano e de suas relações seria negar o próprio humano.

Sob o viés de Morin (2008), o movimento de descomplexificação dos processos complexos, realizado pela ciência clássica, reconhecida como fragmentadora das ciências e dos saberes, dissolveu o que havia de obscuro e não linear com vistas a simplificar suas interpretações. Para tanto, Morin (2008) nominou esse movimento de Paradigma da Simplificação, o que, em síntese, significa a disjunção, a redução e a abstração da complexidade presente no conhecimento científico, de maneira a evidenciar a redução do complexo em simples. Diante disso, a operacionalização cartesiana do conhecimento desintegrou a complexidade em sua microdimensão (o ser individual), assim como distorceu a realidade da macrodimensão do humano (o ser coletivo), gerando, enquanto resultado desse posicionamento, uma inteligência cega (MORIN, 2008).

Nessa ótica, a investigação científica que adota uma visão de sistema fechado, ou seja, sem influências externas a si, oferece apenas “uma visão classificacional, analítica, reducionista, uma causalidade unilinear” (MORIN, 2008, p. 34). A referida

visão fragmentou as ciências e definiu quem tratava do que, tornando-se disciplinar e negligenciadora da inserção de abordagens de cunho inter(trans)disciplinar, ao passo que se recusava à realização desse diálogo entre ciências. Tendo em vista a classificação feita por Morin (2008), este estudo adota a epistemologia do pensamento complexo, nos níveis da microdimensão e da macrodimensão, ao assumir a indissociabilidade do humano na inter-relação com outros sistemas abertos e influenciáveis, como é o caso da língua/linguagem na relação com os aspectos e comportamentos que emergem dessas dimensões em caráter de recursividade.

A fim de corroborar as afirmações apresentadas, expõe-se, de acordo com Morin (2009), três princípios pelos quais a ciência clássica rejeitou a perspectiva complexa: a) o de determinismo universal – tem em vista não só conhecer todo o acontecimento, mas ainda prever todo o acontecimento do futuro; b) o de redução – consiste em conhecer um todo composto a partir do conhecimento dos elementos primeiros que o constituem; c) o de disjunção – adequado a isolar e separar as dificuldades cognitivas umas das outras, o que conduziu à separação entre disciplinas que se tornaram herméticas entre si. As inferências de Morin (2008, 2009) ressaltam a ideia de que o Paradigma da Simplicidade é aquele cujo lema é instaurar ordem onde há desordem, de maneira a expusá-la da descrição e da explicação acerca do objeto de estudo, reduzindo-o a uma ordem fictícia, enquanto lei/princípio. Assim, a simplicidade focaliza o uno e o múltiplo, mas desconsidera o fato de que o uno pode ser, ao mesmo tempo, múltiplo; ela une o que está separado e separa o que está ligado por meio de uma relação dicotômica de disjunção e redução.

No intuito de ilustrar essa posição paradigmática, toma-se a definição de Morin (2008, p. 86) acerca da definição de homem, ou seja, de ser humano, a qual contraria a visão da ciência clássica que, ao “fragmentá-lo” diante de sua constituição, acaba por extrair a complexidade presente e manifestada por ele. Para tanto, Morin (2008) define o ser humano como biológico, logo, dotado de componentes culturais, metabiológicos e inserido em um cenário de linguagem, de ideias e de consciência. Desse ponto de vista, têm-se duas realidades, sendo a primeira, a biológica, e a segunda, a cultural. Mas a ciência clássica esmerou-se para dissociar essas realidades, o que torna o ser humano um objeto/sujeito destituído de aspectos, processos e padrões complexos sob a referida ótica. O Paradigma da Simplicidade entra em divergência com a noção de que o humano biológico está no cultural e vice-versa, dotado, portanto, de múltiplas realidades. Segundo Morin (2008), para sermos nós mesmos, precisamos aprender uma

linguagem, uma cultura, a fim de que possamos nos situar no mundo. Frente ao exposto, cabe lembrar que essa ótica reside em caráter de recursividade, ou seja, que os produtos e os efeitos são causas e produtores do que lhes produzira (MORIN, 2008, p. 108).

Em comunhão a Morin (2008), Nicolescu (1999) defende a necessidade de que as ciências considerem a complexidade e estabeleçam um diálogo de construção de conhecimento que resida no âmbito inter(trans)disciplinar enquanto unificação das ciências. Em “O Manifestado da Transdisciplinaridade”, Nicolescu (1999) discorre sobre a necessidade latente de que as ciências rompam com o caráter fragmentado do fazer científico; do contrário, o isolamento do conhecimento produzido pela extrema especialização de algumas ciências e o não diálogo entre essas não suprirá a urgência de compreensões coerentes relativas aos problemas sociais manifestados. Assim, é no estabelecimento de um diálogo inter(trans)disciplinar que a produção de conhecimento e o fazer científico poderão, possivelmente, evitar turbulências globais com o intuito de desconstruir e ressignificar os construtos epistemológicos e metodológicos outrora sacramentados pelo viés da ciência clássica.

No que é relevante a sua proposta, Nicolescu (1999) define três concepções basilares da transdisciplinaridade, a saber: a) os níveis de realidade – níveis hierárquicos dotados de princípios que não são aplicáveis a todos eles, mas que, na ótica transdisciplinar, são não lineares e interconectados; b) a lógica do terceiro incluído – a concepção do sagrado no científico; c) a complexidade – entendida como abordagem que vem sendo instaurada mais nas Ciências Sociais e Humanas do que nas puras/duras. Em síntese, a transdisciplinaridade visa a que as divergências entre as ciências, na realização de um diálogo inter(trans)disciplinar, sejam o impulso necessário para a compreensão de que os objetos de estudo são complexos e que, portanto, precisam de um fazer científico que seja inter(trans)disciplinar.

Desse modo, na busca pelo estabelecimento da transdisciplinaridade, Nicolescu (1999) enfatiza que “a transdisciplinaridade se interessa pela dinâmica gerada pela ação de vários níveis de realidade ao mesmo tempo” e, ainda, afirma que “as pesquisas disciplinares e transdisciplinares não são antagonistas, mas complementares”. Em outras palavras, essa complementaridade, presente na proposição de Morin (2008) e Nicolescu (1999), consolida a necessidade de um fazer científico, onde a complexidade é inerente aos processos. Nada é mais justo, portanto, do que supor que diferentes perspectivas científicas unificadas possam possibilitar um viés inter(trans)disciplinar, o qual potencializa e ratifica observações complexas. Apresentado, brevemente, o Paradigma

da Simplicidade e a proposta por uma declaração de unificação das ciências, tratar-se-á, a seguir, de apresentar o Interacionismo Sociodiscursivo (ISD), corrente teórica que se configura como ciência do humano, ao integrar diversas ciências em suas postulações, a fim de explorar a concepção de desenvolvimento humano possibilitado pela língua/linguagem em práticas interacionais concretas. Conquanto se está articulando esta pesquisa no nível da recorrência de signos linguísticos, observáveis a partir da técnica de mineração de dados, via ferramenta Sobek, tomaram-se os conceitos de Atividade de linguagem e Ação de linguagem como centrais no recorte teórico-metodológico de pesquisa, para observar, em diálogo com a Teoria da Complexidade, o modo como se pode conjecturar a interpretação de possíveis níveis de complexidade entre o conteúdo de um texto-referência (externo) e um texto-produzido (empírico).

## **2.2 Interacionismo Sociodiscursivo (ISD)**

Neste subcapítulo, apresenta-se a concepção que norteia o presente estudo, que é o desenvolvimento humano pautado *na e pela* atividade de linguagem dos sujeitos e as implicações relacionadas a ela. Logo, a afiliação deste estudo ao ISD deve-se à compreensão de que sua base teórico-epistemológica pode auxiliar na interpretação do recorte feito pela ferramenta de mineração de dados Sobek, a qual, ao extrair dados de textos-empíricos (entidades semióticas altamente complexas), gera um grafo constituído apenas por palavras que, aqui, são consideradas signos linguísticos. Esses signos, quando extraídos de um texto-referência para uma atividade de produção textual, inicialmente, podem ser vistos como envelopes vazios (BRONCKART, 2006, 2008), dotados de total arbitrariedade na relação significante-significado, porém sua recorrência ou não, em caráter de estabilidade ou variabilidade composicional, no texto-produzido pelos sujeitos, pode revelar a complexidade da ação de linguagem e do desenvolvimento humano, por meio de atividades de linguagem, prescritivas de ordem reguladora, em contexto formal de ensino.

Frente ao processo automatizado realizado pela ferramenta, cabe investigar a composição do conteúdo apresentado que, *a priori*, parece evidenciar um caráter regulador nos textos, ou seja, a identificação da emergência de um padrão (recorrência de signos) e/ou comportamento (ação de linguagem), que corresponde ao significado

das representações socialmente construídas e atribuídas aos conceitos, ou seja, evidencia aspectos associados ao nível da atividade de linguagem (coletiva). Porém, faz-se relevante questionar: o que a recorrência ou a inserção de novos conceitos ao texto-empírico representa, no nível da ação de linguagem (individual). E é, justamente, nesse questionamento que o aporte teórico do ISD, em princípio, parece fornecer subsídios para a interpretação dos dados que emergem da transição do processo de leitura de um texto-referência para a elaboração de um texto-empírico, derivado da ação de linguagem dos sujeitos. Neste estudo, os textos-referência (capítulo de livro e artigo científico), assim como os textos-empíricos (textos narrativos-descritivos) estão subentendidos como pertencentes ao gênero secundário (BAKHTIN, 2006), por serem complexos, escritos e situados em contexto de produção de esfera acadêmica, mais especificamente, em turmas de nível de ensino superior, graduação. Ao levar em consideração o escopo apresentado, a fim de apropriar-se dos pressupostos teóricos do ISD, primeiro, apresenta-se brevemente a posição teórica do ISD, seguida da explanação de conceitos basilares dessa teoria, para fins de um pleno entendimento da linha de raciocínio que se quer conduzir neste estudo.

### **2.2.1 A concepção teórica do ISD**

Uma vez que o objeto de estudo das Ciências Humanas é o humano, esta tem como referência a Sociologia, a Psicologia, a Linguística e as ciências da Educação, de onde extrai construtos teórico-epistemológicos com vistas a possibilitar a compreensão de comportamentos, signos e demais aspectos socioculturais relacionados ao humano em desenvolvimento. Dessa maneira, entende-se que é nesse campo do conhecimento que o Interacionismo Sociodiscursivo (ISD) está situado, ou seja, como uma ciência do humano, relacionada à linguagem, que tem como principal fonte de referência as postulações de Saussure (1916), Habermas (1987), Adam (1990), Ricouer (1977), Bakhtin (2003, 2006) e Vygostky (1998), dentre outros autores. Portanto, o ISD, como corrente interacionista-social, fortemente baseada nos construtos de Vygotsky (1998), afirma que é nos processos discursivos caracterizados enquanto “instauração do agir como unidade de análise do funcionamento humano” (BRONCKART, 2008, p. 9) que se constitui a compreensão das características da linguagem e das implicações de seu funcionamento, de onde surgem estratégias de intervenção prática e processos de

mediação. Dessa perspectiva decorre o fato de que o ISD assume três princípios gerais, apresentados, *ipsis litteris*, neste estudo, conforme Bronckart (2006, p. 9), quais sejam:

- a) o problema da construção do pensamento consciente humano, o qual deve ser tratado paralelamente ao da construção do mundo dos fatos sociais e das obras culturais, sendo os processos de socialização e os processos de individuação (ou de formação das pessoas) duas vertentes indissociáveis do mesmo desenvolvimento humano;
- b) o questionamento das Ciências Humanas deve apoiar-se no admirável corpus da filosofia (do espírito de Aristóteles e Marx) e deve, simultaneamente, considerar os problemas de intervenção prática (e principalmente os de intervenção no campo escolar); e
- c) convém contestar a divisão dessas Ciências em múltiplas disciplinas e subdisciplinas, que é decorrente de uma adesão à epistemologia positivista herdada de Comte, tendo em vista que os problemas centrais de uma ciência do humano envolvem, de um lado, as relações de interdependência que se instauram e se desenvolvem entre os aspectos fisiológicos, cognitivos, sociais, culturais, linguísticos etc. do funcionamento humano e, de outro, os processos evolutivos e históricos por meio dos quais essas diferentes dimensões se geraram e se co-construíram.

Diante dos princípios referidos pelo ISD, é possível considerar que o Paradigma da Simplicidade, mencionado no subcapítulo anterior (MORIN, 2008, 2009), assim como a busca por um fazer científico inter(trans)disciplinar (NICOLESCU, 1999) são abarcados pelo ISD, na medida em que esse se propõe a uma compreensão que unifica os princípios das ciências que tratam do humano (Antropologia, Sociologia, Psicologia, Linguística e Educação). De acordo com Bronckart (2008, p. 109), o quadro teórico do ISD comporta um conjunto de princípios que se resume em considerar que o desenvolvimento humano, enquanto evolução em diálogo com princípios do materialismo, monismo e do evolucionismo - ou seja, toda matéria está em constante evolução, assim como os objetos e os processos de pensamento da espécie humana, mesmo que alguns objetos pareçam físicos e outros psíquicos, dada sua constituição fenomológica - ambos são matéria.

No que tange ao evolucionismo, o ISD afirma que, em cada etapa evolutiva, não se pode negar a organização interna de objetos e seres vivos em relação às propriedades de caráter interacional que tais elementos estabelecem com o meio externo. Nesse viés, a evolução humana situa-se em uma perspectiva dialética e histórica, sendo necessário conceber que o processo genealógico humano decorre de um movimento permanente. Essa constatação é corroborada pela noção de que é preciso rejeitar qualquer concepção essencialista do ser humano com vistas a analisar suas capacidades sob um enfoque

genealógico, pois, como afirma Bronckart (2008, p. 111), “não se pode compreender o humano a não ser compreendendo sua construção ou o seu vir-a-ser”. Ainda sobre o escopo do quadro teórico do ISD, afirma-se que ele concebe o caráter interacional dinâmico, adaptativo e complexo, tanto da constituição humana quanto da língua/linguagem (atividade de linguagem). E, neste tocante, Bronckart (2008) assume que o ambiente humano percebido como espaço de práticas constitui-se não somente pelo meio físico, mas, também, pelas condutas dos membros da espécie humana, de modo que ele é organizado por atividades coletivas complexas. A referência à complexidade das atividades humanas remete o posicionamento interacionista sociodiscursivo ao entendimento de que o humano influencia e é influenciado por seu contexto, o que demonstra que o ISD leva em consideração o caráter recursivo pelo qual tudo é organizado e se organiza frente às interações humanas. Portanto, o posicionamento do ISD dialoga com a premissa epistemológica da Teoria da Complexidade que comporta este estudo.

Frente a isso, essas atividades organizam e mediatizam a relação entre os sujeitos e o meio no qual estão inseridos. Assim, atividades não verbais se articulam, segundo Bronckart (2008), às atividades languageiras na forma de acordos, os quais estabelecem condições de produção condizentes com os contextos, com vistas a assegurar-lhes um caráter regulador, como é o caso da atividade de produção textual.

Partindo dessa noção, o ISD tem como objetivo afirmar que as práticas languageiras circunscrevem-se como os instrumentos principais do desenvolvimento humano, no que se refere às capacidades do agir e da constituição identitária dos sujeitos em desenvolvimento. Ainda, nesse pensamento, Bronckart (2006, p. 11) reflete:

Sustentar que as práticas languageiras são o fundamento (isto é a origem) de qualquer funcionamento especificamente humano não nos deve levar a negar a realidade indiscutível de emergência de capacidades cognitivas gerais, que tendem a se liberar das restrições sociais, culturais ou languageiras[...].

Essa afirmação corrobora a perspectiva vygostkiana e considera que a construção de capacidades cognitivas universais deve ser compreendida como resultado de um processo segundo o qual se aplica, progressivamente, as capacidades de pensamento, sendo essas, inicialmente, marcadas pelos aspectos socioculturais e pela linguagem. Em outras palavras, o ISD se dispõe a aperfeiçoar a compreensão das condições e características da atividade de linguagem no quadro do problema do desenvolvimento humano e assim objetiva (BRONCKART, 2006, p. 14):

- i) reconvocar a abordagem vygotskyana e a questionar sua base filosófica;
- ii) reexaminar, com aprofundamento na obra de Saussure, o papel da apropriação de signos na emergência da consciência humana e, por fim,
- iii) estudar os efeitos produzidos pelo domínio dos gêneros de texto e dos tipos de discurso no desenvolvimento, tanto em suas dimensões epistêmicas quanto praxiológicas.

Para Vygotsky (1998), assim como na perspectiva piagetiana, o ISD assume que o modo de funcionamento do pensamento é consciente é de ordem praxiológica. Assim, esse modo é resultante da interiorização das propriedades das atividades coletivas e da linguagem. Baseado nesse entendimento, Bronckart (2006, p. 16) afirma que a linguagem evolui em duas direções funcionais: a social, tendo em vista o rápido desenvolvimento das capacidades de comunicação verbal; e a individual, no que concerne à interiorização dos signos linguísticos (pensamento consciente). A respeito disso, Bronckart (2006) assume a hipótese de que o papel decisivo, como se refere o autor, da interiorização dos signos sociais na constituição e no desenvolvimento do pensamento consciente está presente no papel que a atividade de linguagem desempenha na construção das próprias ações e conhecimentos do sujeito. Referindo-se a essa noção, Bronckart (2006, p. 17) sustenta que é “a interiorização posterior das relações predicativas da língua que permitem que essas unidades se organizem de acordo com o regime de “implicações de significações” que caracteriza as operações do pensamento consciente”. Neste sentido, de acordo com Bronckart (2006), é o discurso que permite que a atividade de linguagem possibilite o desenvolvimento humano na forma de pensamento consciente. Sendo assim, define-se esse conceito por meio de dois posicionamentos:

- a) na acepção dominante, esse termo designa a realidade primeira da linguagem, que é a de ser atividade significante; portanto, a nosso ver, o termo atividade de linguagem (ou languageira) é suficiente para dar conta desse nível e, além disso, duplicar essa expressão com a de “discurso” leva, de fato, a se admitir que há outra realidade mais fundamental que ela (por exemplo o sistema da língua); e
- b) é no nível das unidades linguísticas designadas pelo termo “discurso”, em nossa acepção, que se desenvolvem os processos mais fundamentais de “formatação semiótica” das representações que transitam entre os registros do coletivo e do individual.

Na interpretação dos posicionamentos mencionados, Bronckart (2008) afirma que o interesse da corrente teórica do ISD é o de focalizar o caráter ativo e interativo dos enunciados e discursos, de modo a entender a natureza da atividade de linguagem relacionada aos processos desenvolvimentais humanos. Nessa perspectiva, a linguagem é concebida pelo ISD como a competência humana para criar organizações semióticas, ou seja, significantes, capazes de assegurar a comunicação e o pensamento dos sujeitos socialmente organizados, como instrumentos do agir. Cabe ressaltar que, em diálogo com a abordagem da Teoria da Complexidade, apresentada a seguir, neste estudo, o ISD considera que o agir e a atividade humana, concretizados por meio de práticas discursivas e dado o grau de interdependência dessas, mostram dinamismo permanente, característica essa presente na língua/linguagem enquanto fenomenológica.

Bronckart (2008, p. 16) afirma que “a linguagem só existe em práticas, e essas práticas, ou jogos de linguagem, são heterogêneas, diversas e estão em permanente transformação”, de modo que essa é correlativa a das formas de vida, no que se refere ao agir humano. O ISD ainda define que práticas languageiras, conceito central neste estudo, são os instrumentos de regulação do agir geral e, na relação com esse agir, surgem as unidades e estruturas mobilizadas assumindo suas significações (BRONCKART, 2008). Infere-se, então, que é da contextualização de uma atividade ou ação de linguagem que se pode entendê-la como produto do uso e/ou resultado de uma prática languageira significativa, elaboradora de conhecimentos humanos constitutivos das representações oriundas do conjunto do aparelho cognitivo humano, conforme afirma Bronckart (2008).

Diante do exposto, é possível considerar que o ISD compartilha do mesmo posicionamento manifestado pelos autores que abordam a Teoria da Complexidade, ou seja, que tanto a língua/linguagem e o ser humano se enquadram na definição de Sistemas Adaptativos Complexos (SACs), uma vez que o estado inicial de conhecimento dos sujeitos só pode ser arrolado como um sistema aberto, pois é, ao considerar os comportamentos observáveis dos sujeitos e seu discurso, que se pode interpretar o conjunto de suas transformações. Bronckart (2008, p. 19) ainda ressalta a noção de que “a ação não pode ser objeto de uma explicação causal”, pois ela precisa ser compreendida por meio de um processo interpretativo. No que tange à noção de transformação, ou seja, o caráter evolutivo de uma língua/linguagem, cabe elencar três princípios aos quais o ISD se opõe radicalmente, que se originam da tradição filosófico-

científica que, inquestionavelmente, ainda orienta a maioria das abordagens das Ciências Naturais e Humanas/Socias.

O primeiro princípio é o da estabilidade e da finitude dos mecanismos de organização do universo, tendo esses por fundamentação a Física newtoniana, a qual diz que todos os processos atestáveis apresentam um caráter determinista e reversível (os elementos seriam finitos e as transformações observáveis resultantes de propriedades estáveis e eternas). Corroborando, mais uma vez, para com a abordagem da Teoria da Complexidade, Bronckart (2006) faz questão de ressaltar que tal abordagem reducionista acaba por desconsiderar o caráter dinâmico e imprevisível da evolução, tendo sido apontado pela noção de Termodinâmica, quando da emergência de novos padrões ou comportamentos em escala evolutiva sistêmica e contínua.

O segundo princípio refere-se ao fato de que não se pode obter total controle e previsibilidade sobre qualquer objeto de estudo - aceitação da legitimidade do recorte dos objetos e subobjetos de conhecimento -, de maneira que Bronckart (2006, p.125) infere que tal noção cartesiana de controle absoluto das variáveis que constituem a produção de um conhecimento científico é, na verdade, um equívoco para as ciências que nela se apoiam. Diante disso, o ISD comunga da posição dinâmica e histórica herdada de Darwin ou de Marx e afirma que as ciências que se negam a tal perspectiva evolutiva aceitaram vetar a abordagem das relações de interdepência entre os aspectos psicológicos, cognitivos, sociais e culturais relacionadas ao desenvolvimento humano. Este princípio opõe-se, radicalmente, ao paradigma simplicista (MORIN, 2008), pois considera que os sistemas estão inter-relacionados e não apresentam sempre os mesmo padrões e/ou comportamentos humanos.

O terceiro princípio é o de que os fenômenos físicos e psíquicos são dissociáveis, de modo que o ISD assume a noção da unidade de substância material e da continuidade dos processos, dos quais emerge o psiquismo humano sobre as transformações das formações sociais e dos fatos psicológicos. Portanto, o ISD aceita o caráter dialógico da interação social desenvolvimental, inscrita em um auditório social de múltiplos sistemas em caráter de recursividade. Tendo em vista a breve apresentação do escopo que fundamenta o ISD, a seguir, mostram-se conceitos centrais que norteiam esta teoria proposta por Bronckart (1999, 2006, 2008).

### 2.2.2 Signo linguístico, atividade de linguagem e ação de linguagem

Na medida em que Bronckart (1999, 2006, 2008, 2007) toma emprestado e reformula conceitos propostos pelas correntes teóricas, que dão subsídios para que o ISD possa conjecturar-se com ciência do humano, considerou-se necessário, antes de tratar da noção de apropriação de signos linguísticos e como esta reverbera em uma ação fundante do pensamento consciente, delimitar alguns conceitos basilares, já na acepção do ISD, tais como: agir, atividade, ação, atividade de linguagem e ação de linguagem (BALTAR, 2007). Diante dessa demanda, cabe considerar que um conceito-chave no ISD é o de agir, o qual se deriva da noção de ação humana e designa toda e qualquer forma de intervenção orientada, individual ou coletivamente, dentro de práticas humanas realizadas dentro de formações de grupos sociais. O termo *atividade*, por sua vez, relaciona-se à interpretação do agir, que demanda a mobilização de dimensões motivacionais e intencionais, no nível coletivo, para compor formações sociais. Já o termo *ação*, em sentido restrito, tem as mesmas demandas, porém está situado no nível da ação individual dos sujeitos, que é relativo a suas condutas.

No que diz respeito ao construto *atividade de linguagem* (discurso), entende-se que ele é o responsável pelo fenômeno coletivo de produção e circulação de textos que, segundo Bronckart (1999), tem por objetivo os modos de interpretação dos contextos interacionais e das atividades humanas. A atividade de linguagem está entendida, segundo Bronckart (2006, p.104), como “o estatuto de mecanismo de ajuste elaborado pelos humanos para planificar, regular e avaliar as atividades coletivas”. No que tange ao conceito de *ação de linguagem*, esse compõe o caráter dialógico e polifônico da atividade de linguagem (coletiva), mas ele representa as formas de agentividade de um ator único (sujeito) acerca dos processos semióticos que ele desencadeia ao aceitar, rejeitar e/ou reformular a validade social da língua/linguagem. Apesar disso, é preciso ter em mente o que é da ordem do coletivo (atividade linguagem) e o que é relativo à ordem do indivíduo, como forma de não se equivocar quanto ao que diz respeito ao que é socioculturalmente produzido e validado (atividade de linguagem – gêneros textuais (gêneros do discurso – BAKHTIN, 2003, 2006), e tipos de discurso – e ao que é individualmente produzido e semiotizado (textos-empíricos), muito embora ambas as ações estejam implicadas em uma relação circular de apropriação da linguagem.

Ainda na delimitação terminológica proposta por Bronckart (2006, 2008), cabe considerar que o sujeito/pessoa que exerce alguma prática de ação é intitulado por atuante/*actante*, ao passo que, no nível interpretativo, adota-se o termo *ator*, uma vez que o actante é parte constitutiva do processo interpretativo, possibilitado por suas capacidades, motivos e intenções, e o termo *agente* é designado quando o atuante é fonte do processo (BALTAR, 2007). Embora, neste estudo, não sejam explorados os conceitos de gêneros textuais, tipos de discurso e atitudes enunciativa e discursiva, esses são tomados como dados, em se tratando de atividades que envolvem textos externos à ação dos atuantes, bem como em relação a textos-empíricos. Portanto, a concepção de texto para o ISD se aproxima da acepção bakhtiniana de enunciado/texto/discurso, sendo este entendido como

“unidade comunicativa verbal, oral ou escrita, gerada por uma ação de linguagem, que se acumula historicamente no mundo das obras humanas, que os indivíduos utilizam para interagir uns com os outros nos diferentes ambientes discursivos da sociedade” (BALTAR, 2008, p.147)

Na medida em que essas atividades humanas (interação) se efetuam na forma de textos-discursos, os quais são unidades semióticas e comunicativas textualizadas, infere-se que se tornam mais ou menos adaptadas aos tipos de atividades e tipos de interações humanas. Assim, esses textos encontram-se, na ótica de Bakhtin (2003) e de Bronckart (1999), como gêneros indexados em arquitexto inserido em uma determinada comunidade. Para fins de observação desses textos, Bronckart (2006) afirma que o ISD constrói sua análise didática/escolar mediativa, visando a analisar as produções das aquisições coletivas e do desenvolvimento dos sujeitos, quando esses as integram e as transformam a partir de aspectos presentes nas atividades de linguagem, uma vez que podem representar transformações, que são dinâmicas e permanentes, no que se refere à constituição dos sujeitos (atuantes), enquanto organismo humano pensante e consciente.

Nessa perspectiva, as transformações situam-se no nível individual e coletivo em relação a construtos pré-construídos por um coletivo/social antecedente às práticas dos sujeitos. Para tanto, a interpretação dessa transformação tem por base a teoria saussuriana, baseada em três aspectos, quais sejam: a) o caráter social da língua – atividade coletiva humana e seu caráter histórico, ou seja, uma língua sofre transformações ao longo do tempo, dadas as forças sociais que a conduzem a isso; b) o estatuto do pensamento puro, independente ou previamente a sua formatação *em e por*

signos; c) a necessidade de compreender as dimensões físicas e psíquicas em relação ao papel determinante da atividade de linguagem ou discursiva. Segundo Bronckart (2006), as ações humanas podem ser observadas sob dois ângulos, o sociológico e o psicológico. O sociológico abarca a condição atestável do fluxo contínuo das ações, nas quais diversos sujeitos estão inseridos em formações sociais diversas, ao passo que, do ponto de vista psicológico, a constituição dessas formações sociais acerca da ação humana demanda promover um corte nos fluxos das ações sociais atestáveis para que se possa analisar a ação e o enquadramento de um sujeito singular nessa configuração. Uma vez feito o corte, é possível avaliar a conduta exercida pelo sujeito, em caráter individualizado, em relação às intenções e motivações dos sujeitos.

Para o ISD, a referida atividade de linguagem (coletiva), composta por uma rede de conhecimentos, é objeto de avaliação dos sujeitos que a compõem e está associada a três mundos: a) objetivo - avaliações das pretensões à verdade; b) social - pretensões à adequação às normas; c) subjetivo - pretensões à veracidade. Nesse sentido, a atividade de linguagem constitui a atividade social e dos mundos formais que configuram o contexto em que os sujeitos interagem. Em outras palavras, ao integrar uma formação social, o sujeito constrói suas representações particulares, validando-as em relação às coordenadas dos três mundos, sendo então uma racionalidade secundária à racionalidade social, praticada por uma atividade coletiva. Portanto, é preciso integrar uma formação social coletiva de base sociocultural para que o sujeito valide, coletivamente, representações intrapessoais e, em seguida, possa criar representações interpessoais semiotizadas que, por vezes, são arbitrárias às motivações e intenções do coletivo.

Relativo a isso, infere-se que é no coletivo que as bases representacionais da atividade de linguagem são construídas, ou seja, adquirem forma e conteúdo semanticamente construídos, mas é na ação de linguagem assumida pelo sujeito, individualmente, que essa funda a constituição semiotizada da validade representacional das significações do conteúdo (signo linguístico). Assim, a relação das representações interpessoais decorre das representações intrapessoais, posto que uma primeira mediação entre sujeito e mundos está alicerçada na atividade de linguagem como processo interpretativo objetivo, social e subjetivo. As representações coletivas podem ser compreendidas como mecanismos de criação de unidades semiológicas socioconvencionais, como infere Bronckart (2008, p.24), ao considerar estas postulações de Habermas (1987): a) arbitrárias por não se basearem em propriedades

naturais e objetivas das entidades que exprimem; b) sociais por estabilizarem as relações entre entidades verbais e seus referentes, pois determinam um acordo ou convenção entre membros de determinado grupo.

O signo linguístico, depositário das representações coletivas, na perspectiva saussureana, cristaliza-se na medida em que os sujeitos, coletivamente, atribuem a sua produção sonora representações do mundo, adquirindo, assim, valor representativo ou declarativo, porém, em caráter dialético, “nunca estão definitivamente estabilizados, pois evoluem permanentemente” (BRONCKART, 2008, p.27). Para Bakhtin (2006), todo signo é ideológico, sendo tal ideologia um resultado/reflexo do social, uma transformação social que, em perspectiva sistêmica, desencadeia uma modificação da língua também. A transformação, em caráter evolutivo, é regida por leis internas e, sobretudo, externas, ou seja, de cunho social. De acordo com Bronckart (2008), a primeira tese do ISD defende que a atividade de linguagem é produtora de objetos de sentidos ao atribuir valor semântico (no nível coletivo) e semiótico (no nível individual); já a segunda tese refere-se ao fato de que, por ser uma atividade social, a atividade de linguagem compõe um pensamento humano que é semiótico e social ao mesmo tempo.

Ao partir do pensamento bakhtiniano, Bronckart (2008) reafirma que “signos-ideias não podem provir da atividade de indivíduos isolados; ao contrário, são, necessariamente, resultado dos discursos produzidos no quadro das interações sociais” o que lhes confere caráter dialógico. Os signos são, portanto, “entidades arbitrárias e desdobráveis (formas coletivas que reduplicam imagens individuais), e essas características fazem com que os signos possam “refletir” ou expressar qualquer realidade” (BRONCKART, 1999, p.169). Por sua vez, Bakhtin (2006) afirma que o signo é mutável diante de processos enunciativos cercados por elementos de base ideológica em uma situação social específica, de modo que afeta a significação. O signo, então, caracteriza-se por possuir uma natureza orgânica e móvel.

Para Bakhtin (2006), a palavra (signo) pode ser compreendida como um indicador das mudanças decorrentes do contexto social interacional, assumindo sua posição de que a língua, na interação dialética, apresenta a empiria das mudanças em um processo criativo e ininterrupto (BAKHTIN, 2006, p.64). Dessa forma, Bakhtin (2006) entende que todo signo está sujeito aos critérios de avaliação ideológica (isto é: se é verdadeiro, falso, correto, justificado, bom, etc.), sendo ele mais do um reflexo ou sombra da realidade, pois é um fragmento da realidade do sujeito, ou, como menciona o

autor, uma encarnação material. Para o autor, “a própria consciência só pode surgir e se afirmar como realidade mediante a encarnação material em signos” (BAKHTIN, 2006, p. 32). Ressalta-se, também, a afirmação de Bakhtin (2006, p.48) de que “psiquismo interior não deve ser analisado como uma coisa; ele não pode ser compreendido e analisado senão como um signo”. Referindo-se às mudanças no sistema de um determinado signo, Bakhtin (2006) assegura que a mudança/transformação de um elemento do sistema do signo, conseqüentemente, gera um novo sistema. Essa inferência, percebe-se, é plenamente compatível com a noção de interdependência dos SACs. É nessa perspectiva que esta pesquisa se situa, ou seja, nas transformações dos fluxos que interferem na elaboração de um determinado sistema, neste caso, o texto-empírico.

Prosseguindo no aprofundamento do ISD, compreende-se, frente às postulações apresentadas, a necessidade de abordar as questões relacionadas à dinâmica da atividade humana de produção textual, sendo essa central nesta tese, com vistas a investigar qual é a relação existente entre a atividade de linguagem e a ação de linguagem, quando da utilização de uma ferramenta digital que extrai palavras (signo) com o intuito de orientar o processamento cognitivo dos sujeitos que dela fazem uso. Além dos conceitos referidos, no próximo subcapítulo, apresenta-se a Teoria da Complexidade, mais especificamente quanto à descrição de Sistemas Adaptativos Complexos (SACs), objetivando que, ao longo do fazer-científico deste estudo, a atividade de linguagem (assim como a ação de linguagem) seja vista como um sistema que é adaptativo, complexo, posto que a linguagem se configura como elemento influenciador e influenciado das ações humanas. Dessa forma, atividade e ação de linguagem estão averbadas no fluxo das ações humanas como SACs, que estão perpetuamente interconectados e possibilitam a identificação de possíveis níveis de complexidade do conteúdo minerado, ligados às zonas de desenvolvimento humano (VYGOSTKY, 1998), em relação à manifestação deste ou daquele signo.

Retomemos a noção de produção textual e suas implicações para que seja feita uma aproximação precisa do objeto de estudo desta tese. Cabe ressaltar que, embora o processo inferencial de leitura (decodificação e codificação) (KOCH e ELIAS, 2012; MARCUSCHI, 2008) esteja implicado na transição da ação de linguagem do atuante para a produção de um texto empírico, essa não é aprofundada neste estudo, uma vez que o que está em questão é conteúdo pós-leitura, ou seja, os signos extraídos pela ferramenta digital, de modo automatizado.

Quanto à produção de um texto, por mais que esteja atrelada a outros modelos (gêneros textuais), ela sempre se constituirá como um texto singular, ou seja, empírico, pois apresenta uma unidade concreta de produção de linguagem, sob a condição de um gênero, contendo tipos de discurso, mas mais do que isso, ela é a manifestação das tomadas de posição individuais de seu autor/produtor em relação à determinada atividade de linguagem, neste estudo, referente à produção de textos-empíricos escritos. As manifestações postas em prática e organizadas em relação ao conteúdo temático de um texto-empírico, originam-se e variam de acordo com o nível de desenvolvimento do autor/produtor, de modo que elas sejam armazenadas em sua memória, o que lhe permite exercer sua ação de linguagem, tendo em vista noções associadas à arquitetura textual de textos orais e escritos com os quais já teve contato (BRONCKART, 1999).

Dessa forma, acessar experiências prévias quanto a textos socialmente construídos significa mobilizar e perturbar sua estrutura lógica, assim como as sequências que lhe compõem; por isso, “nunca podem constituir-se como cópias idênticas das macroestruturas semânticas” (BRONCKART, 1999, p. 98). Os textos empíricos não se configuram como cópias por terem sido semiotizados pelo autor-produtor, ação que exige do sujeito que acesse mundos discursivos de sua constituição identitária, cujas coordenadas, segundo o ISD, são sempre distintas das do mundo ordinário de quem primeiramente o escreveu, sendo esse um texto externo à ação de linguagem do atuante (neste caso, de quem interpreta forma e conteúdo). Por isso, diz-se que quem produz um segundo texto (texto-empírico) toma emprestado e como referência outros textos (gêneros-modelo) e tem liberdade (agentividade) para adaptá-lo às suas necessidades comunicativas, associando a sua produção valores particulares.

Então, todo texto-empírico comporta seu estilo próprio e individual, o que, para a Teoria da Complexidade, abordagem convergente ao ISD e presente nesta tese, equivaleria ao mesmo que dizer que um SAC (aqui a ação de linguagem), ao receber algum tipo de perturbação em sua estrutura, conserva algumas de suas propriedades e altera outras. Nesse horizonte, os conhecimentos humanos e as representações de mundo são elaborados na interação entre a ação individual e o escopo dialógico da língua/linguagem e, “mesmo quando são alvo de uma reorganização singular, resultante da dimensão experiencial própria de cada pessoa, continuam portando os traços dessa alteridade constitutiva” (BRONCKART, 1999, p.321).

Segundo Bronckart (2008), os textos podem se distribuir em múltiplos gêneros, os quais são indexados, ou seja, validados ou reconstruídos em situações concretas de

comunicação. Portanto, um texto empírico indexado contém figuras interpretáveis de ação de linguagem, o que permite afirmar que, ao interpretá-las se está analisando a ação humana. Relativo a isso, compreende-se que a ação de linguagem, como ação significativa humana, na forma de texto-empírico escrito, é uma obra aberta, pois, como afirma Bronckart (2008, p.34), baseado na tese de Ricouer (1986), essa, após ser produzida “se desliga do agente, desenvolvendo suas próprias consequências, assim como um texto, que produzido e colocado à disposição dos interpretantes, se desliga das intenções de seu autor”. Desse modo, é ao desligar-se de seu autor que os signos são reconfigurados e semiotizados a partir da perspectiva de quem os analisa/lê.

Em síntese, o ISD considera que a atividade de produção textual é uma atividade humana prescritiva e potencializadora da produção de conhecimento, neste viés, em contexto formal de ensino e aprendizagem. Na dinâmica da atividade de produção textual, os sujeitos mobilizam recursos em relação ao contexto, conteúdo temático do texto e relativo às coordenadas associadas aos mundos objetivo, social e subjetivo, pretendendo sua validação a verdades socialmente elaboradas. Em relação à atividade de produção textual, é preciso que os sujeitos manifestem suas capacidades de linguagem, a saber: a) capacidades de ação, b) capacidades discursivas, c) capacidades linguístico-discursivas (CRISTOVÃO, 2007), como referidas por Bronckart (2006, 2008) e citadas ao longo deste subcapítulo. Desse modo, é necessário conceber o contexto de produção, o destinatário, o objetivo do conteúdo, o gênero textual e as regras de funcionamento da linguagem socialmente aceita. Na etapa seguinte, o sujeito precisa pôr em prática suas capacidades discursivas, a fim de que escolha uma infraestrutura geral para seu texto, dotada de um determinado tipo de discurso e sequências textuais, visando, como infere Cristovão (2007, p.263), “a escolha e elaboração de conteúdos, que surgem como efeito de um texto já existente e estímulo para que outro seja produzido”. Em seguida, o sujeito precisa colocar em prática suas capacidades linguístico-discursivas com a finalidade de realizar estas operações: a) textualização (coesão nominal e verbal e coerência); b) mecanismos enunciativos de gerenciamento de vozes e modalizações; c) construção de enunciados; d) escolha de itens lexicais (CRISTOVÃO, 2007, p.263). Como mencionado, este estudo não analisa a totalidade composicional de textos-empíricos, mas a implicação das ações dos sujeitos quanto à escolha e à articulação de signos-linguísticos, na transição do conteúdo de um texto-referência para um texto-produzido. Então, é relevante considerar a contribuição

do ISD acerca das postulações de Saussure (1916), no que concerne à gênese do pensamento consciente.

Acerca disso, o ISD considera que a contribuição saussureana se dá em primeiro nível por considerar o caráter histórico, social e em permanente transformação da língua, seguido da concepção de que essa existe na interioridade dos sujeitos falantes em relação ao coletivo. Além disso, é basilar levar em consideração o papel determinante da semiose e sua ação fundante do pensamento consciente relacionado ao estatuto dos signos linguísticos. A partir dessa premissa, o signo linguístico é, “de início arbitrário, no sentido banal do termo, ou seja, imotivado: a escolha de um significante é independente das propriedades “naturais” do referente ao qual se endereça um significado” (BRONCKART, 2006, p.111). Esse, portanto, é radicalmente arbitrário por sua configuração na relação de formatação entre significante-significado não ser natural, mas composto por um conjunto de imagens mentais, socialmente elaboradas.

Por sua vez, o ISD afirma que o signo linguístico é de caráter imotivado - dada a ausência de fundamento natural, possibilitando autonomia para o estabelecimento de representações permanentes; discreto - por delimitar o funcionamento representativo (estabiliza unidades) instaurando um sistema de operações ou de pensamento (BRONCKART, 2006, p.113); totalmente arbitrário - por potencializar um desdobramento do funcionamento psíquico dos sujeitos relativo às representações sociais. Na medida em que o resultado da análise do recorte automatizado pela ferramenta Sobek, sob o viés do ISD e da Teoria da Complexidade, pode desencadear possíveis formas de mediação da ação de linguagem dos sujeitos, embora este estudo não apresente práticas intervencionistas, a concepção de mediação formativa é, a seguir, descrita na compreensão de que se possa associar os resultados a processos de mediação, seja em contexto formal de educação ou não, em se tratando do procedimento de mineração de dados.

### **2.2.3 Da produção de textos-empíricos à mediação formativa**

O conceito de mediação formativa está abarcado por este estudo com uma postulação do ISD que pode, após a análise de dados, ser retomada, com vistas a sugerir formas de aprimoramento da atividade de linguagem (no nível coletivo), bem como da ação de linguagem (no nível individual) dos sujeitos, no intuito de potencializar práticas

de produção textual e as implicações relativas à utilização de procedimentos de mineração de dados.

A primeira tese do ISD quanto ao conceito de mediação formativa é a de que os processos de mediação estão inscritos como os meios pelos quais os adultos integram os “recém-chegados” ao conjunto das concepções sociais pré-construídas, as quais estão instauradas na atividade de linguagem. Esta concepção arrola sujeitos em geral, não limitando a noção de “recém-chegados” a crianças. Como afirma Bonckart (2006, p. 129), “esse campo de análise diz respeito ao conjunto de processos de controle e avaliação das condutas verbais e não verbais que são implementados desde o nascimento”, referindo-se, também, aos processos educativos explícitos que se realizam, em geral, em instituições escolares, no caso deste estudo, de educação superior – graduação. Ressalta-se, contudo, que, neste trabalho, o termo avaliação refere-se à apropriação de um conhecimento linguageiro, não se tratando, portanto, de um caráter avaliativo pedagógico de desempenho relacionado à atribuição de conceitos/notas, pelo menos, não no presente recorte.

A mediação formativa como descrita pelo ISD, por sua vez, visa a compreender os efeitos de determinada atividade sobre os indivíduos e divide-se em duas abordagens: a) as condições de transformação do psiquismo sensório-motor, fundador da pessoa, enquanto resultado da interiorização das propriedades estruturais e funcionais dos signos linguísticos, conforme a proposta de Saussure; b) as condições de desenvolvimento dos sujeitos e suas capacidades ativas, concernente às representações individuais e coletivas vinculadas aos pré-construtos sociais historicamente situados.

A concepção apresentada entende que a atividade de linguagem e os mecanismos de produção e interpretação que dela emergem, contribuem para a transformação da constituição identitária do sujeito. No que se refere às atividades coletivas, infere-se, segundo o ISD, que elas se organizam por meio de estruturas de cooperação e de colaboração, que são responsáveis e organizadoras não só do pensamento humano consciente, mas, também, dos contextos interacionais. Portanto, a mediação formativa linguageira está subentendida no recorte investigativo deste estudo, como a maneira pela qual o um sujeito (adulto/tutor/professor) orienta outros sujeitos (crianças/adultos/ alunos) a transpor dificuldades na transição de um estágio de conhecimento (zonas do conhecimento). Essa concepção está pautada nos construtos de Vygotsky (1998) e, também, é revisitada em sua formatação original, no subcapítulo 2.4

desta tese, quando da aproximação das postulações deste autor à Teoria da Complexidade.

Retomemos esta conceituação, vinculada às condutas humanas em relação ao agir dos sujeitos, considerando que o ISD define dois níveis distintos: a) o nível coletivo das atividades orientadas e b) o nível singular das ações. O primeiro é aquele que diz respeito às finalidades e aos motivos de um grupo, e o segundo refere-se às “porções” (destaque do autor) das atividades coletivas de responsabilidade individual de um sujeito, o qual está imbuído de representações de suas próprias intenções, razões e capacidades. Em consonância com a proposta deste estudo, o ISD afirma que a observância da atividade de linguagem, por meio da identificação das dimensões intencionais e motivacionais das condutas individuais, só pode ser apreendida no curso das condutas e não previstas de antemão, perspectiva essa que evidencia a complexidade inerente ao uso da linguagem. Bronckart (2006, p. 242) define o agir como “um mecanismo de pilotagem das condutas em redes de determinações complexas e entrecruzadas”, o que faz com que o ISD assuma a Teoria do Agir, a partir de Habermas (1987). A referência à pilotagem diz respeito à inserção da atividade de linguagem aos pré-construtos sociais que a antecedem, ou seja, o modo pelo qual são relevantes ou não, no que concerne ao uso da linguagem enquanto pretensão de validação. Esse posicionamento faz com que o ISD (BRONCKART, 2006, p.244) defenda que há uma determinação pessoal nas interações co-construídas:

É essa determinação pessoal que confere aos indivíduos uma “liberdade” e uma “possibilidade de criatividade” em relação aos pré-construídos coletivos, que lhes permite avaliá-las, transformá-las, reorientá-las e, enfim, que permite que não sejam apenas “agentivizados” pelo social, mas que sejam também atores da vida social ou construtores de fatos sociais.

Ainda referente à mediação formativa *na e pela* atividade de linguagem, o sujeito mediador é encarado como agente e pessoa, pois ele está inserido em um contexto no qual medeia processos languageiros, sendo definido como pessoa no mundo, o que faz com que possa intervir mais ou menos, dependendo de sua intencionalidade quanto ao processo formativo dos sujeitos mediados. Concernente a contextos educacionais, o ISD caracteriza a ação docente como mediadora de processos de construção do conhecimento e ressalta que essa tem papel fundamental no andamento (VYGOTSKY, 1998) daqueles que necessitam de algum tipo de

mediação/intervenção para alcançarem outros níveis de domínio do conhecimento para além dos já consolidados acerca de determinada ação ou assunto. Nesse cenário, o mediador assume a responsabilidade de prover intervenções que suscitem no outro a transformação de seu comportamento, visando a que esse assuma sua responsabilidade de sujeito pessoa e agente, podendo, assim, avaliar, transformar, reorientar sua atividade de linguagem/ação de linguagem.

Diante do entendimento de que as postulações oriundas do ISD foram elecandas para a composição desta pesquisa, com a cautela de não adentrar perspectivas outras senão aquelas que cabem a este estudo, uma vez que não se está operando com modelos prescritivos propriamente ditos, mas com a construção de um conhecimento empírico que emerge da própria análise do objeto de estudo. Desse modo, a seguir, discorre-se acerca da Teoria da complexidade, de maneira a, inicialmente, tratar do que se compreende por pensamento complexo e, posteriormente a isso, refletir sobre sua instauração no campo da Linguística Aplicada, levando em consideração a pertinência dessa para definir a atividade de linguagem e Ação de Linguagem como Sistemas Adaptativos Complexos (SACs). Enfatiza-se que a adesão à Linguística Aplicada, ciência que tem por função o estudo da linguagem e seus desdobramentos, deve-se ao fato de ela poder auxiliar na compreensão da dinamicidade da linguagem, trazendo conhecimentos específicos desse campo para a área da Informática na Educação quanto à prática de mineração de textual.

### **2.3 A Teoria da Complexidade na Linguística Aplicada**

Uma vez tendo percorrido a respeito da necessidade de aderir a uma teoria que abarque a complexidade presente no fazer científico, apresentar-se-á o que, de fato, se compreende por complexidade, bem como se fará a descrição da Teoria da Complexidade situada no campo da Linguística Aplicada como ciência que está para além do fazer científico reducionista constitutivo da ciência clássica. Desse modo, então, questiona-se: afinal de contas, o que é, de fato, a complexidade e quais são suas características? De acordo com Morin (2008, p. 20):

à primeira vista, a complexidade é um tecido (*complexus*: o que é tecido em conjunto) de constituintes heterogêneos inseparavelmente associados: coloca o paradoxo do uno e do múltiplo.

Nesta perspectiva, Morin (2008) afirma que, na relação entre constituintes, enquanto tecido de acontecimentos, estão as ações, as interações, as retroações, as determinações e as bifurcações que são passíveis de observações desses fenômenos complexos, tal como se refere Bronckart (2006, 2008) aos conceitos de atividade e ação de linguagem. Referente a isso, neste estudo, abarca-se o caráter fenomenológico que reside na relação sujeito (práticas humanas), linguagem e tecnologia, no que diz respeito à aprendizagem e ao desenvolvimento humano, mediado e potencializado por recursos tecnológicos (LÉVY, 1999). Sob este viés, ao considerar que a linguagem é fenomenológica, enquanto discurso ou texto-discurso (BRONCKART, 2006; 2008) não se pode conjecturar que essa não seja diacrônica e sincrônica ao mesmo tempo e se constitua em um auditório social, pois seu estabelecimento emerge *no e pelo* seu uso. E, de acordo com Larsen-Freeman e Cameron (2008), o discurso (atividade de linguagem) é uma atividade complexa e dinâmica de uso da linguagem (ação de linguagem).

Portanto, por menor que seja a manifestação de um padrão ou mudança no âmago dos processos linguísticos, seja em caráter oral ou escrito (como no caso desta pesquisa), essa sempre é relevante, com vistas a visualizar seus efeitos e causas, em caráter de recursividade, as quais em um contexto de aprendizagem, por exemplo, podem receber o devido tratamento (mediação), a fim de atingir possíveis metas, previamente estipuladas por um sujeito/tutor/professor/mediador. Quanto a isso, infere-se que se trata de prospectar e não tomar por certo que o que é oferecido é aprendido, seja por intermédio da linguagem ou qualquer outro meio instrumento que vise à aprendizagem de um conteúdo ou ação (LARSEN-FREEMAN e CAMERON, 2008).

Então, retomando-se a definição de complexidade, encontra-se em Morin (2008, p. 151) a afirmação de que a ela é “a dialógica entre ordem/desordem/organização, de maneira que, por trás da complexidade, ordem e desordem dissolvem-se, sendo que a complexidade traz à tona a metafísica da complexidade”. Neste compasso, Capra (2002) e Morin (2008) sugerem que, quando uma tessitura complexa é afetada por um determinado fluxo entrópico (energia externa – neste estudo a ação de linguagem semiotizada do atuante), essa, mesmo longe de um estado equilibrado, devido à interação de seus componentes com o meio, pode supor um estado novo de modos de pensar, saber e fazer, de onde emergem novas estruturas e formas de ordem.

A desordem não significa total falta de ordem, pelo contrário, ela é reveladora da ordem que está imbricada aos processos caóticos e desorganizados, seja quanto ao uso da linguagem, a um fenômeno climático, a um comportamento humano, etc. Ainda

quanto à ciência clássica, referida anteriormente, acredita-se que o pensamento complexo “não substitui a separabilidade pela inseparabilidade – ele convoca uma dialógica que utiliza o separável, mas o insere na inseparabilidade” (MORIN, 2000, p. 200). Apresentada, brevemente, a base científico-filosófica acerca da complexidade no campo filosófico-educacional, quanto ao desenvolvimento do humano, doravante, assume-se o que fora noutra momento apenas denominado como tessitura complexa, como Sistema Adaptativo Complexo (SAC).

Nesse entendimento, o biólogo e autor da Teoria Geral dos Sistemas, Bertalanffy (1973, p.84), infere que “um sistema complexo pode ser definido como um complexo de elementos em interação”. Em posse desta demoninação, Morin (2008, p.43) acrescenta alguns aspectos, dizendo que, no sistema complexo, tem-se “uma relação entre partes que podem ser muito diferentes umas das outras e que constituem um todo que é, simultaneamente, organizado, organizando e organizador”. E é nesse sentido que se observa a relação entre o que emerge de um processo de mineração, na relação entre atividade e ação de linguagem quanto aos signos linguísticos manifestados pelos sujeitos na transição do conteúdo do texto-referência para o texto-produzido. No que diz respeito a esse construto teórico, Bertalanffy (1973), assim como, posteriormente, professaram Larsen-Freeman e Cameron (2008), reforçou a necessidade de avançar na instauração de uma perspectiva sistêmica nas Ciências Sociais, enquanto ciência do humano, como proposto pelo ISD.

Defronte a essa necessidade e com o propósito de avançar no fazer científico inter(trans)disciplinar, situa-se o estabelecimento do pensamento complexo (MORIN, 2008) em relação ao humano e à linguagem, já na Linguística Aplicada, porém sem a pretensão de apresentar um estado da arte desse tema.

A perspectiva inter(trans)disciplinar possibilitada pelo pensamento complexo na Linguística Aplicada, ocorreu por meio da adesão à Teoria da Complexidade, em 1997, quando a linguísta Diane Larsen-Freeman, pesquisadora norte-americana da Universidade do Michigan, tornou-se uma das precursoras da inserção epistemológica dos Sistemas Adaptativos Complexos (SACs) aos estudos linguísticos no campo do desenvolvimento humano e da língua/linguagem. Então, ao publicar o artigo “*Chaos/complexity science and second language acquisition*”, na *Applied Linguistics*, em 1997, a autora despertou a curiosidade de outros pesquisadores que, aparentemente, cansados de suas abordagens teóricas simplicistas, em certa medida, se interessaram pelo tema. De acordo com Corrêa (2012), Larsen-Freeman (1997) realizou a árdua

tarefa de possibilitar um panorama geral das características dos SACs, visando ao devido ajuste terminológico e ao registro de inúmeros trabalhos de pesquisadores que, próximos ou já na perspectiva da Teoria da Complexidade, mesmo que não assumidamente, vinham desenvolvendo pesquisas em caráter inter(trans)disciplinar, no campo da língua/linguagem, sob este viés.

Na medida em que Larsen-Freeman (1997) situa os termos oriundos das Ciências Biológicas, Químicas e Físicas ao campo das ciências humanas, a autora estabelece um diálogo inicial que acabou por receber amplo acolhimento. No que diz respeito à Teoria da Complexidade em relação ao conceito de SACs, Larsen-Freeman (1997, p.147) afirma que “os estudos de sistemas complexos não lineares, e o estudo da língua/linguagem têm muito em comum”<sup>2</sup>. Frente a isso, pesquisadores, no mundo e no Brasil, assumiram o compromisso de se dedicarem a um fazer científico de cunho inter(trans)disciplinar, cuja motivação parece ser a de incorporar ciência e filosofia, visando a produzir conhecimento acerca de processos fenomenológicos, tendo por base a Teoria dos Sistemas Complexos, a Teoria do Caos e a Teoria da Auto-Organização, as quais oferecem subsídios científicos para a designação do termo Teoria da Complexidade. Embora essa ainda não tenha recebido o devido crédito em algumas ciências, como afirma Nicolescu (1999) e Larsen-Freeman e Cameron (2008), a ela tem sido amplamente adotada nas ciências que colocam o desenvolvimento humano no centro de suas investigações, como se propõe o ISD.

Aprofundando seu movimento inicial de 1997, Larsen-Freeman, publicou, em 2008, a obra *Complex Systems in Applied Linguistics*, em parceria com sua colega Lynne Cameron, a qual foi, sem dúvida, bem aceita, como se pode notar pelas referências feitas às autoras, em artigos e livros, por linguístas, no exterior e no Brasil. Posteriormente, tem-se a publicação de *Language is a Complex Adaptive System: Position Paper*, de autoria de Beckner *et alii* (2009), onde Larsen-Freeman e Lynne Cameron fazem-se presentes pelo movimento inter(trans)disciplinar de divulgação e instauração da Teoria da Complexidade na Linguística Aplicada.

O fazer científico da Linguística Aplicada brasileira também manifestou interesse pela utilização da Teoria da Complexidade, o que pode ser observado nas publicações de muitos linguístas (CORRÊA, 2012; FRAGA, 2004, 2007a, 2007b;

---

<sup>2</sup> Do original: “*The study of complex nonlinear systems and the study of language have much in common*” (LARSEN-FREEMAN, 1997, p. 147).

LEFFA, 2009; MARTINS, 2008; PAIVA, 2002, 2005, 2006; PARREIRAS, 2005; SADE, 2009; SILVA, 2008; SOUZA, 2011; VETROMILLE-CASTRO, 2007; dentre outros). De acordo com Corrêa (2012), Larsen-Freeman e Cameron (2008) mencionam alguns temas de interesse da Linguística Aplicada em que diferentes abordagens têm sido realizadas, a saber: o modelo dinâmico para descrever o desenvolvimento e a perda de vocabulário; a aquisição de segunda língua; a mudança na proficiência de multilinguagem em nível individual para a dinâmica do multilinguagem; a evolução da língua; a metáfora no discurso, mostrando como a língua e o desenvolvimento conceitual interagem, etc.

As áreas de estudo mencionadas corroboram a proposta de Larsen-Freeman (1997) de que o desenvolvimento linguístico do indivíduo seja compreendido a partir de pensar o complexo em suas interações viabilizadas *na* e *pela* linguagem. Estabelecida a repercussão da Teoria da Complexidade e do pensar o complexo na Linguística Aplicada, compreende-se que observar e explicar um processo demanda conceber uma gama complexa de componentes que constituem um SAC. Então, no que confere ao humano o estatuto de SAC, diz-se que ele é indissociável, por exemplo, ao contexto social, sendo esse um dos fatores mais relevantes em sua constituição, segundo autores como Bakhtin (2003; 2006), Vygotsky (1998), Bronckart (2006, 2008), Larsen-Freeman (1997) e Larsen-Freeman e Cameron (2008), os quais são adotados em diferentes momentos deste estudo.

Nessa perspectiva, é na interação do humano com outros sistemas e subsistemas que está a linguagem, cuja função é permitir a comunicação entre os sujeitos, de modo que essa não pode ser compreendida fora do âmbito social. A linguagem, por sua vez, exerce um papel fundamental na sociedade humana, assim como na cultura, permitindo que o conhecimento cultural, gerado e transformado, por essa, possa ser transmitido, elaborado e ressignificado em um movimento contínuo de apropriação de saberes, nos mais variados contextos (BECKNER *et al.*, 2009). O que se quer afirmar é que cérebro, cognição, consciência, interação humana, sociedade, cultura e história estão todos inter-relacionados, em diversos níveis complexos e dinâmicos, à língua/linguagem.

A Teoria da Complexidade, portanto, não dissocia o social, o físico e o cognitivo, pois esses estão em constante interação e mudança. Os três elementos relacionados à atividade humana são sistemas abertos, ou seja, são influenciados uns pelos outros. Para tanto, Larsen-Freeman (1997) afirma que a Teoria da Complexidade

trata do estudo de sistemas complexos, dinâmicos, não lineares, auto-organizáveis, abertos, emergentes e, às vezes, caóticos e adaptativos.

Parece haver um questionamento latente quanto ao se afirmar que o humano e a linguagem estão sob investigação neste estudo. Porém, de imediato, essa inquietação é elucidada a partir de duas afirmações na ótica da complexidade e dos SACs: a primeira diz respeito a que todas as vezes que a linguagem é usada ela muda (LARSEN-FREEMAN, 1997) e a segunda, “se mudo, o mundo muda...” (LEFFA, 2009). Então a mutabilidade, nessa lógica, afeta quem produz a mudança, assim como o que fora produzido em caráter espiralado, atingindo diferentes SACs inter-relacionados. Portanto, esse é um movimento complexo, não linear, caótico, adaptativo e emergente, no qual o humano e a linguagem são indissociáveis. Conquanto que o ISD reconhece a atividade de linguagem e a ação de linguagem na perspectiva dos SACs, a seguir, com o objetivo de aprofundar suas características, em uma perspectiva mais epistemológica interpretativista, são descritas suas definições no sentido de serem aplicadas à análise da ação de linguagem submetida ao recorte feito pela ferramenta Sobek, no que tange à relação entre o conteúdo do texto-referência e o do texto-produzido. Contudo, ressalta-se que a tomada de posição para a apresentação desses conceitos levou em consideração apresentá-los, quase que em sua configuração original, ou seja, nos termos das ciências que lhes formalizaram, de modo que possam, até mesmo, ser adotados por outras abordagens científicas que não somente as de cunho educacional e/ou linguístico.

### **2.3.1 Complexo, dinâmico, adaptativo e não linear**

A definição de sistema complexo (LARSEN-FREEMAN, 1997; LARSEN-FREEMAN e CAMERON, 2008; MORIN, 2008) é a de qualquer sistema, enquanto “organismo vivo” que sofra influências externas e seja formado por agentes/componentes/fatores/ (partes) que compõem a perspectiva de um todo, os quais, conectados entre si, interagem e mudam sua forma em diferentes escalas de tempo, gerando, a partir dessa interação, novas estruturas (sociais, físicas ou cognitivas). Essas, por sua vez, emergem do comportamento de seus agentes que, ao apresentarem mudanças em caráter heterogêneo, não são totalmente previsíveis frente a sua imprevisibilidade. Seria, portanto, uma impossibilidade teórica a adoção de um único princípio teórico descritivo-explicativo para diferentes níveis/comportamentos

complexos. O sistema complexo descrito é, pois, dinâmico na medida em que “tudo muda” e nada em sua estrutura é completamente estático. Dessa maneira, os próprios elementos, agentes e/ou processos em um sistema complexo podem ser sistemas complexos. Disso, entende-se que “um sistema de atividade humana ou de desenvolvimento será dinâmico em cada nível de organização social ou humana, seja no nível sociocultural, no individual, no neural ou no celular (LARSEN-FREEMAN e CAMERON, 2008, p.29)<sup>3</sup>. Em outras palavras, o sistema complexo comporta um ecossistema (WALDROP, 1992; MORIN, 2008) dotado de uma estrutura complexa (desequilíbrio – organização – reorganização – equilíbrio) que é um todo dinâmico, ordenado em organização e organizador, na relação com seus agentes e vice-versa (MORIN, 2008).

A adaptação às turbulências às quais o sistema se submete ou é submetido, neste caso, em se tratando do humano em contexto de aprendizagem, exige que ele assuma o caráter de adaptativo, pois a adaptação torna-se um dos aspectos que permite que o sistema não fique estático diante das mudanças internas e externas. Então, entre uma turbulência/fase/nível/zona e outra, os agentes atingem níveis de experiência, ou seja, certa estabilidade ou domínio, poder-se-ia dizer, uma consciência em nível superior, servindo para que reavaliem constantemente suas condições para atuarem naquele e em outros níveis de organização (GLEICK, 1989).

Resultante da dinamicidade que constitui o sistema, origina-se a não linearidade, que ocorre devido à interação descentralizada entre os agentes e os componentes interconectados de modo que diferentes níveis e agentes influenciam, de forma não hierarquizada, o sistema como um todo. Causa e efeito, portanto, não são proporcionais, ou seja, são não lineares. Dessa forma, a interação pode conduzir níveis e agentes para diferentes direções, de modo que, as mudanças tendem a ser desproporcionais. Por exemplo, a ampliação do vocabulário de sujeitos de uma comunidade de fala não acontece proporcionalmente, pois está atrelada às interações pertinentes que eles realizam. Percebe-se, contudo, que a não linearidade na aprendizagem é cíclica, adaptativa e dinâmica, pois a interiorização mais estável<sup>4</sup> de domínios do conhecimento,

---

<sup>3</sup> Do original: “*A system of human activity or development will be dynamic at each level of social or human organization, from the sociocultural, through the individual and down to the neural and celular*”. (LARSEN-FREEMAN e CAMERON, 2008, p.29).

<sup>4</sup> De acordo com Larsen-Freeman e Cameron (2008, p. 33), “... estabilidade não significa estagnação, mas representa a dinâmica do sistema que mantém sua identidade geral, sem ser sujeito às flutuações agressivas ou mudanças caóticas”. (Do original: “... stability does not mean stagnation and stasis, but

do comportamento e as formas do pensamento podem ser notadas não só a longo prazo, mas em diferentes níveis do processo. Na próxima subseção apresenta-se o que se compreende por caótico, imprevisível e sensível às condições iniciais em situação de mudança sistêmica.

### **2.3.2 Caótico, imprevisível, sensível às condições iniciais**

Um sistema complexo é caótico ou está à beira do caos por apresentar estabilidade e variabilidade quanto às interações aparentemente desordenadas e aleatórias (LORENZ, 1993) que realiza, ou seja, essas parecem totalmente aleatórias, mas não o são, de fato. Em um curto espaço de tempo, é possível determinar comportamentos previsíveis, porém eles não podem ser submetidos a esse prognóstico a médio e longo a prazos. Embora o sistema seja determinístico, sua não linearidade e dinamicidade o tornam caótico devido às imprevisibilidades que emergem das possíveis e constantes interações entre níveis, agentes e subsistemas internos e externos.

Dessa forma, uma pequena alteração (GLEICK, 1989) pode gerar consequências gigantescamente positivas ou negativas no desenvolvimento do sistema, aspecto esse que não pode ser previsto de antemão em sua totalidade. Isso significa dizer que ele é sensível às condições iniciais, ao passo que uma pequena mudança em seu estágio atual de desenvolvimento e aprendizagem, quanto aos fatores e agentes que o cercam e compõem, pode desencadear um fenômeno/comportamento totalmente diferente.

Segundo Larsen-Freeman e Cameron (2008, p.58), “um sistema próximo ou à beira do caos muda adaptativamente para manter sua estabilidade, demonstrando um alto nível de flexibilidade e responsividade”.<sup>5</sup> Mas, de fato, são as turbulências e os atratores que influenciam o sistema e permitem que esse passe por estados de fase que, à beira ou próximo de um estado caótico. As mudanças suaves ou radicais em um sistema, frente às suas transformações, são chamadas de mudança de estado/fase ou bifurcação.

---

*represents a dynamic system that maintains its overall identity, without being subject to wild fluctuations or chaotic change”.)*

<sup>5</sup> Do original: “A system at or near the edge of chaos change adaptatively to maintain stability, demonstrating a high level of flexibility and responsiveness”. (LARSEN-FREEMAN e CAMERON, 2008, p.58).

Nas palavras de Larsen-Freeman e Cameron (2008, p.49) “o espaço de fase ou estado de fase representa um cenário de possibilidades de um sistema e como ele muda e se adapta com o tempo, o sistema se move através desse panorama”<sup>6</sup>. O sistema, ao apresentar esse panorama situacional de suas condições, oferece a visualização de possíveis estados/zonas que poderiam ser alcançados movendo-se de uma zona atual/real para uma possível/prospectada, contudo imprevisível quanto a ser, de fato, atingida.

Assim, no que tange à aprendizagem como um componente em eterno movimento, pode-se dizer que interagir é aprender, pois o sistema pode adquirir novas estruturas/zonas/funções e, posteriormente, transmiti-las a outros sistemas em interação consigo. Portanto, é diante das interações e das turbulências pelas quais o sistema passa que ele potencializa e é potencializado, apresentando a emergência de seus processos/comportamentos. A seguir, apresenta-se o que se entende por aberto, sensível ao *feedback* em relação ao movimento de recursividade apresentado por um SAC.

### **2.3.3 Aberto, sensível ao *feedback***

Um sistema complexo é aberto na medida em que permite que energia/matéria/insumo<sup>7</sup> oriundos de componentes/agentes/contextos externos a si o invadam, gerando mudanças em sua estrutura que se adaptam ao novo contexto interacional por meio da experiência que extraem das influências sofridas. Essa condição de estar aberto faz com que o sistema complexo esteja em desequilíbrio, porém reequilibrando-se após sofrer energia exterior a si, em caráter de auto-organização. Essa auto-organização enquandra-se em um padrão de rede (CAPRA, 2002), o qual se organiza pela recursividade, ou seja, um processo/estado/efeito que evidencia e se alimenta de seus padrões iniciais.

Segundo Prigogine (1997) e Larsen-Freeman e Cameron (2008), os sistemas que se estabilizam por intermédio das constantes mudanças são, às vezes, intitulados de estruturas dissipativas em não equilíbrio. Contudo, cabe ressaltar que sistemas em

---

<sup>6</sup> Do original: “A state space or phase space represents the “landscape of possibilities” of a system, and, as it changes and adapts over time, the system moves through this landscape”. (LARSEN-FREEMAN e CAMERON, 2008, p.49).

<sup>7</sup> Termo cunhado por Rudolf Clausius (1822-1888), físico e matemático, o qual está ligado à Segunda Lei da Termodinâmica e se refere à troca de calor/energia, que é dissipada na interação e acaba por integrar-se a outro corpo, servindo de subsídio para que esse, por sua vez, “alimente-se” e tenha condições de interagir, operando em outras condições que não aquelas iniciais.

equilíbrio estão em constante adaptação às mudanças contextuais, em contraste aos sistemas em não equilíbrio que, além de se adaptarem aos contextos, acabam por neles causar mudanças. Dessa maneira, os sistemas complexos adaptativos e abertos são sensíveis às condições iniciais e aos resultados das turbulências que atravessam, o que deixa evidente que o tipo de *feedback* (LARSEN-FREEMAN e CAMERON, 2008, p. 84) que recebem assume um papel revelador em relação às novas estruturas que emergem do processo adaptativo. Na sequência dessa conceituação, descreve-se o que são atratores e fractais em situações de autossimilaridade quanto aos agentes e componentes de um SAC.

### 2.3.4 Atratores estranhos e fractais

Um atrator é uma região do estado de fase/espço do sistema pelo qual ele (o sistema) tende a se mover (LARSEN-FREEMAN e CAMERON, 2008, p. 50). O surgimento de um atrator deve-se à energia/insumo que é trazida do exterior de sua estrutura. Tratam-se, portanto, de atratores estranhos por se diferenciarem dos componentes já estabelecidos em seu interior. É relevante ter em mente que um atrator propicia um novo posicionamento, movendo o sistema por meio da emergência de novas orientações que, ao o afetarem, exigem que ele se adapte e, em relação às suas condições iniciais, mantenha-as em certa medida, mas que, ao final, poderão estar quase que irreconhecíveis.

As mutações sofridas em seu turno apresentam autossimilaridades que podem ser reconhecidas mediante uma investigação minuciosa, uma microscopia<sup>8</sup> do todo. Assim, os atratores podem incitar ordem em um determinado estado de fase/espço e não necessariamente em todos os níveis interacionais dos componentes, o que revela que uma perspectiva fractalizada, por vezes, faz-se necessária, a fim de desvelar regiões específicas atingidas pelos atratores. O conceito de fractal (MANDELBROT, 1982) descreve padrões autossimilares, porém com infinitas subdivisões internas, o que pode revelar variabilidade mutacional em certos fractais,<sup>9</sup> e não, em outros. Dessa forma, o

---

<sup>8</sup> A observação de objetos e seus pormenores com dimensões inferiores ao limite da resolução do olho humano. Em outras palavras, uma perspectiva atenta aos detalhes.

<sup>9</sup> Para Gleick (1989), a ótica fractalista é a realização de um recorte no processo, com o intuito de investigar certa ramificação ou comportamento geradores de alterações sistêmicas em escala menor ou

grau de variabilidade das influências causadas pelo atrator torna-se um fator relevante para verificar o quão estável está o sistema em um momento específico.

Quando o sistema é afetado por um atrator, ele tende a manter-se estável, reorganizando-se em sua estrutura interna, mas, se diante de forças externas “fortes”, o sistema tende a desequilibrar-se, necessitando reorganizar-se, considerando as propriedades que o afetaram, isso lhe conduzirá à emergência sistêmica de novas formas. Podem-se elencar três tipos de atratores (LARSEN-FREEMAN e CAMERON, 2008, p.56): atratores de pontos fixos, atratores cíclicos e atratores caóticos. Os atratores de pontos fixos são os mais simples, pois representam o movimento do sistema para um estado preferível e estável, permanecendo assim por um período de tempo maior. Os atratores cíclicos são aqueles em que o sistema se move periodicamente entre diferentes e variados estados. Os atratores caóticos ou atratores estranhos<sup>10</sup> caracterizam-se como uma região ou estado de fase/espço em que o comportamento do sistema se torna instável e frágil, pois, ao menor sinal de uma perturbação, o sistema se metamorfoseia de um estado para outro, assumindo um estado caótico, ou seja, nominado como um atrator caótico (LARSEN-FREEMAN e CAMERON, 2008, p.57). Já o atrator caótico traz à tona uma possível mudança de fase/nível/comportamento, mas determinar as mudanças que ocorrerão de maneira precisa não é possível, pois haveria a necessidade de se ter controle total das alterações que estão acontecendo no interior do sistema, para prospectar sua condição de possível emergência. Por fim, na seguinte subseção, são descritos os conceitos de auto-organização e emergência.

### **2.3.5 Auto-organização/Emergência**

Após passar por uma mudança de fase/estado, um sistema complexo tende a apresentar auto-organização ou emergência, o que significa que essa mudança fez com que ele evidenciasse propriedades dinâmicas que o guiaram às novas estruturas internamente organizadas. Após a mudança, inquestionavelmente, sua estrutura é

---

maior. Por outro lado, um fractal mesmo que diretamente relacionado às condições iniciais de um SAC não pode revelar seu todo estrutural.

<sup>10</sup> Termo cunhado pelos matemáticos Floris Takens e David Ruelle.

alterada, mas isso não significa que todas os seus componentes/níveis/agentes tenham sido alterados. Logo, esse fator revela que, embora aparentemente diferente de sua constituição inicial, o sistema ainda manteve níveis/escalas/estados inalterados. Segundo Larsen-Freeman e Cameron (2008, p.59), “a emergência enfatiza o surgimento de um novo estado em um nível de organização superior àqueles prévios. O comportamento ou fenômeno emergente tem algumas totalidades reconhecíveis”.<sup>11</sup>

Em acordo com Larsen-Freeman e Cameron (2008), Morin (2008, p.105), quanto ao conceito de emergência, infere que

A emergência é um produto da organização que, embora inseparável do sistema enquanto todo, aparece não só ao nível global, mas eventualmente também ao nível dos componentes. Assim, as qualidades inerentes às partes no seio dum dado sistema estão ausentes ou virtuais quando estas partes se encontram isoladas; só podem ser adquiridas ou desenvolver-se pelo e no todo [...].

No que concerne, por exemplo, aos conceitos aprendizagem e desenvolvimento, cabe evidenciar que, embora o fenômeno emergente permita um panorama de novas estruturas estáveis (por vezes, emergentes de uma desordem organizacional prévia), essas continuarão abertas às potencialidades de novas forças externas ao sistema. Portanto, isso reforça a noção de que o surgimento de futuras emergências é algo imprevisível, pois diferentes graus de variabilidade, flexibilidade e responsividade continuarão afetando o sistema, uma vez que esse é um sistema adaptativo, complexo, dinâmico e auto-organizável.

De acordo com Larsen-Freeman e Cameron (2008, p.61), a auto-organização ou emergência pode revelar que “os resultados da auto-organização e emergência podem ser negativos, neutros ou positivos em termos humanos”<sup>12</sup>. Os resultados - positivos ou negativos - são, de todo modo, relevantes, pois revelam uma relação de causa e efeito, a qual evidencia as razões pelas quais certos fatores são alterados ou não. Larsen-Freeman e Cameron (2008, p. 75) reforçam essa constatação mediante um contexto de aprendizagem, inferindo que “os professores podem planejar suas lições

---

<sup>11</sup> Do original: “*Emergence is the appearance in a complex system of a new state at a level of higher organization than the previous one. The emergence behavior or phenomenon has some recognizable wholeness*”. (LARSEN-FREEMAN e CAMERON, 2008, p.59)

<sup>12</sup> Do original: “*The outcomes of self-organization and emergence may be negative, neutral, or positive in human terms*” (LARSEN-FREEMAN; CAMERON, 2008, p. 61).

cuidadosamente, mas nunca podem saber o que os aprendizes trarão para as lições planejadas, assim os resultados são imprevisíveis”<sup>13</sup>.

Ainda, quanto à interação realizada, o sistema complexo pode ressaltar tanto propriedades que são de nível mais alto, assim como de nível mais baixo, mas que, todavia, estão interconectadas. Um exemplo coerente de que não se trata da necessidade de uma força externa de alto nível para afetar a estrutura do sistema por meio de uma atividade de baixo nível (pouco complexa) é o citado por Larsen-Freeman e Cameron (2008, p. 59): “a emergência no aprendizado ocorre quando novas ideias enquadram-se em um momento de *insight*. Uma vez compreendidas, o novo conhecimento influencia as outras ideias.”<sup>14</sup>

Em se tratando de uso da linguagem, poder-se-ia dizer que é da interação entre os sujeitos que padrões e comportamentos linguísticos emergem, adaptando-se às diferentes situações e coadaptando-se aos recursos com os quais interagem. Dessa forma, ao passo que os sujeitos adequam o uso da linguagem aos novos contextos em que se inserem, acabam por perpetuar um movimento contínuo de mudança. Do contrário, seria pertinente dizer que as línguas e as múltiplas formas de linguagem entrariam em extinção, como algumas entraram, se não possuíssem como SACs os sujeitos dotados de interação verbal. A língua emerge da atividade de linguagem, primariamente, para comunicar e, *a posteriori*, para estabelecer formas de atividade (LARSEN-FREEMAN e CAMERON, 2008). Levando em consideração a teorização realizada em relação aos SACs, expõe-se uma síntese de suas características, segundo Larsen-Freeman e Cameron (2008, pp. 69-70):

- sistemas complexos estão em constante mudança;
- sistemas complexos podem mudar suave ou radicalmente;
- sistemas complexos podem apresentar auto-organização por meio de mudança de fases que produzem comportamento ou fenômeno emergente em níveis mais altos de organização;
- sistemas complexos *em* e *próximos* de atratores apresentam ambos estabilidade e variabilidade;
- sistemas complexos próximos à beira do caos são muito flexíveis e responsivos;

---

<sup>13</sup> Do original: “*Teachers can plan lessons carefully but can never exactly know what learners will bring to the plan lesson, and so outcomes are unpredictable*”. (LARSEN-FREEMAN; CAMERON, 2008, p. 75).

<sup>14</sup> Do original: “*Emergence in learning occurs when new ideas fall into place in an a –ha moment*”. (LARSEN-FREEMAN; CAMERON, 2008, p.59).

- sistemas em regiões caóticas de seus estados de espaço (sistemas caóticos) são extremamente sensíveis às pequenas mudanças em relação às condições iniciais;
- alguns sistemas complexos apresentam criticalidade próximos de fases de transição;
- dois ou mais sistemas complexos podem mudar em resposta um ao outro, em um processo de coadaptação;
- sistemas complexos podem ser descritos como variáveis coletivas, e suas mudanças de fase por meio de parâmetros de controle.<sup>15</sup>

Em se tratando de contextos formais de interação e tendo em vista as definições apresentadas, entende-se que os seres humanos também se enquadram na definição de SACs. Portanto, um mediador/tutor/professor tem a função de observar a natureza complexa da ação de linguagem, visando a prover atividades que possam afetá-los a ponto de que o que é prospectado para uma atividade possa ser alcançado como resultado de uma emergência da ação humana. Nessa perspectiva, uma tarefa de aula assume em si a definição de trajetória/processo, ou seja, um sistema complexo, dentre outros, co-construído pelos sujeitos que compõem turma/grupo (LARSEN-FREEMAN e CAMERON, 2008; LEFFA, 2009). Alguns alunos, ao co-construírem interações com outros alunos e também com o professor, poderão manifestar emergências sistêmicas em contraste a outros que não manifestarão no mesmo nível. Isso exigirá que o mediador/tutor/professor encontre formas de afetá-los com vista ao “domínio” de um aspecto necessário para tornar o sistema emergente. Quanto a isso, compreende-se que nenhum conteúdo/plano de aula é único e fixo, pois ele deve emergir das decisões entre alunos e professores, muito embora, na maioria dos casos, ele seja elaborado previamente, o que não quer dizer que tenha que ser mantido como inalterável.

Embora essa perspectiva não esteja em questão neste estudo, cabe levar em consideração que as variações coletivas, características de uma turma, podem relevar as causas e efeitos da emergência da produção de conhecimento e possíveis ações futuras em relação a zonas de desenvolvimento dos alunos, no nível coletivo, ao passo que, no

---

<sup>15</sup> Do original: “Complex systems are always changing; complex systems can change smoothly or radically; complex systems can show self-organization through phase shift that produce emergent behavior or phenomena on a higher level of organization; complex systems in and near attractors show both stability and variability; complex systems near the edge of chaos are very flexible and responsive; systems in chaotic regions of their state space (chaotic systems) are extremely sensitive to small changes in the initial conditions; some complex systems show criticality around phase transitions; two or more complex systems can change in response to each other in a process of co-adaptation, and; complex systems can be described with collective variables, and their phase shifts through control parameters”. (LARSEN-FREEMAN; CAMERON, 2008, pp. 69-70).

nível individual, isso demanda uma abordagem fractalizada, como postula Bronckart (1999, 2006, 2008), um corte no fluxo das ações coletivas, a fim de verificar a constituição da aprendizagem e do desenvolvimento de um determinado sujeito/aluno. A produção de conhecimento por intermédio de práticas discursivas, orais ou escritas, está relacionada aos acoplamentos que os sujeitos fazem, uns às realidades dos outros, pelo uso da ação de linguagem, pois acoplar-se é adquirir e adaptar-se às informações de outro sujeito. Nessa interação, os sujeitos estão constantemente equilibrando, desequilibrando e reequilibrando estruturas de conhecimento pré-estabelecidas, o que reforça a ideia de que “o conhecimento é um processo” (LARSEN-FREEMAN e CAMERON, 2008, p.131)<sup>16</sup>. Referente à noção de texto, essa é considerada como produto da linguagem em uso, caracterizando-se como atrator caótico e emergente dentro da organização da trajetória de SAC, seja ele o humano ou a ação de linguagem, sendo o produto (discurso ou texto) e processo (interação/processo de produção), ambos relevantes para a compreensão do desenvolvimento humano físico, cognitivo e social.

Ao se considerar a relevância do aporte teórico de Vygotsky (1998) nesta pesquisa, por meio de uma aproximação à Teoria da complexidade, mais especificamente sobre zonas de desenvolvimento, acredita-se que tanto o aprendizado quanto o desenvolvimento são ativados por meio da inter-relação entre os ecossistemas mentais internos e os ecossistemas sociais externos ao sujeito. Com isso, afirma-se que, em toda interação entre partes, há aprendizagem e desenvolvimento, embora se diga que quem aprende é o sujeito menos proficiente em determinado tema; na realidade, o sujeito que realiza o andamento (VYGOTSKY, 1998) do outro também acaba por adquirir elementos dessa experiência, seja em âmbito formal ou informal. Dito isso, retoma-se, a afirmação de Larsen-Freeman e Cameron (2008, p.185) relativa à definição de discurso ou texto (escrito), entendendo-se que esse, por si só, não apresenta dinamicidade, pois é fixo e imutável, e os processos que o tornam parte integrante de múltiplos sistemas dinâmicos são a leitura, a escrita e o uso.

O processo de leitura pode alterar o panorama (percepção cultural e processo cognitivo) do sujeito (SAC), quando da aceitação ou rejeição de determinado tema, por exemplo, com isso afetando seu espaço de fase, o que pode levar o sujeito à adaptação de suas concepções prévias ao texto. Já o exemplo da leitura exige que os sujeitos, cognitivamente, estabeleçam correlações com experiências prévias, alfabetização

---

<sup>16</sup> Do original: “*Knowledge is a process*”. (LARSEN-FREEMAN e CAMERON, 2008, p.131).

(letramento), outros textos lidos e assim por diante, de modo a encontrar significado para o texto. Nesse tocante, diferentes sujeitos podem traçar trajetórias diversas a partir da leitura de um mesmo texto. Segundo Larsen-Freeman e Cameron (2008, p.187), os atratores inseridos no processo de leitura podem ser compreendidos com as interpretações estáveis, podendo haver variabilidade nas interpretações, “onde as partes do texto estiverem abertas a interpretações múltiplas”<sup>17</sup>.

No que é relativo ao processo de escrita, esse pode ser descrito como um sistema complexo em si, pois, em geral, manifesta a totalidade das características outrora mencionadas, o que permite afirmar que o texto emerge de uma gama de sistemas e subsistemas (sistema motor ao cognitivo, indo da escrita de versões, mudanças e adaptações rumo a uma versão final). A versão final de um texto, portanto, é pertencente a um determinado gênero textual (BRONCKART, 2008) e é vista como a emergência do processo de leitura e escrita, pois daí passará a ser utilizada para fins imprevisíveis. Na próxima seção, propõe-se a convergência dos princípios vygostkianos à Teoria da Complexidade.

## **2.4 Os princípios de Vygotsky à luz da Teoria da Complexidade**

Neste subcapítulo, realiza-se uma aproximação de conceitos centrais postulados por Vygotsky (1998), próxima à perspectiva de Silva (2009), por meio de citações diretas que se considerou revelante apresentar integralmente, uma vez que o objetivo é estabelecer a relação dessas com os pressupostos teórico-epistemológicos da Teoria da Complexidade. Embora essa estratégia acompanhe a linha de raciocínio de Silva (2009)<sup>18</sup>, ela visa a postular que as zonas de desenvolvimento são, na verdade, estados de fase, mais do que SACs. Portanto, não se trata de uma reconstrução, mas da identificação da notável aproximação dos pressupostos vygostkianos à luz da Teoria da Complexidade, aqui, em diálogo com Morin (2008), Bronckart (2006) e Larsen-Freeman e Cameron (2008). Então, ao levar em consideração que é na interação com o

---

<sup>17</sup> Do original: “...where parts of the text are open to multiple interpretations”. (LARSEN-FREEMAN e CAMERON, 2008, P.187)

<sup>18</sup> Esse subcapítulo segue a ilustre abordagem conduzida por Silva (2009).

outro que se concretiza a apropriação de possíveis novos modos de pensar, saber e fazer, cabe enfatizar que o conceito de mediação formativa, apresentado anteriormente, aqui está elencado, inicialmente, como a mediação entre pares, ou seja, o auxílio que um sujeito oferece ao outro, mesmo que de forma não consciente, para a realização de uma determinada atividade (leitura, produção textual, intervenções pedagógicas, etc). Inicialmente, as inquietações quanto aos conceitos de aprendizado e desenvolvimento fazem com que Vygotsky (1998, p.87) afirme que a relação entre esses conceitos é confusa quanto às teorias do desenvolvimento, dizendo que

a relação entre aprendizado e desenvolvimento permanece, do ponto de vista metodológico, obscura, uma vez que pesquisas concretas sobre o problema dessa relação fundamental incorporam postulados, premissas e soluções exóticas, teoricamente vagos, não avaliados criticamente e, algumas vezes internamente contraditórios: disso resultou, obviamente, uma série de erros.

Diante dessa obscuridade, Vygotsky (1998) resume tal problemática a partir de três grandes posições teóricas (p.87), as quais são, brevemente apresentadas:

a) A primeira centrada no fato de que os processos de desenvolvimento da criança são independentes do aprendizado, sendo este externo ao desenvolvimento. Nessa linha de raciocínio o aprendizado era dependente dos avanços relacionados ao desenvolvimento, de modo que esse antecede o aprendizado não desempenhando um papel significativo a respeito do desenvolvimento ou na maturação de funções ativas durante o processo de aprendizagem. Assim, segundo Vygotsky (1998, p.89) “o desenvolvimento ou a maturação é visto como pré-condição do aprendizado, mas nunca como resultado dele.

b) A segunda posição teórica, frente à primeira, postulava que o aprendizado é o desenvolvimento em si, sendo que “o desenvolvimento é visto como o domínio dos reflexos condicionados, não importando se o que se considera é o ler, o escrever, ou a aritmética, isto é, o processo de aprendizado está completa e inseparavelmente mistura com o processo de desenvolvimento” Vygotsky (1998, p.89).

c) A terceira abordagem tentava superar os extremos das duas anteriores, de modo a combiná-las, propondo que o desenvolvimento se estruturava em dois processos diferentes, mas se influenciavam, onde de um lado a noção de maturação estava ligada diretamente ao desenvolvimento do sistema nervoso e, do outro lado, o aprendizado era por excelência um processo de desenvolvimento Vygotsky (1998, p. 90-91).

Tendo conhecimento dos três posicionamentos teóricos apresentados, Vygotsky (1998) rejeita-os e assume diferentemente das correntes psicológicas que o antecederam:

a compreensão de desenvolvimento não poderia estar unicamente atribuída aos fatores internos observáveis diante de testes de estímulo-resposta, quanto aos processos cognitivos, sem considerar os processos sociais, ou seja, o contexto interacional. Vygotsky (1998, p. 45) sustenta seu posicionamento:

Dominados pela noção de mudança evolucionária, a maior parte dos pesquisadores em psicologia da criança ignora aqueles pontos de viragem, aquelas mudanças convulsivas e revolucionárias que são tão frequentes no desenvolvimento da criança. Para a mente ingênua, evolução e revolução parecem incompatíveis, e o desenvolvimento histórico só está ocorrendo enquanto segue uma linha reta. Onde ocorrem distúrbios, onde a trama histórica é rompida, a mente ingênua vê somente catástrofe, interrupção e descontinuidade. Parece que a história para de repente, até que retome, uma vez mais, a via direta e linear de desenvolvimento.

Essa abordagem está arrolada por Bronckart (2006) que, em parte, se apropria das postulações de Vygotsky (1998) para evidenciar a relevância da interação social e a importância dos fatores socialmente pré-construídos antes mesmo que a criança frequente a escola, tal como o fez esse autor, o que reforça a compreensão de que, de fato, há uma relação indissociável e complexa entre aprendizado e desenvolvimento. Desse modo, para Vygotsky (1998) e Bronckart (2006), a apropriação do pensamento humano consciente em relação aos signos é um processo não linear. Tendo constatado essa noção, Vygotsky (1998) postulou que o desenvolvimento deriva de um nível social e *a posteriori* do nível individual. Essa máxima retoma o entendimento de que os sujeitos transformam processos interpessoais em intrapessoais, ou seja, a apropriação de aspectos pré-construídos pode conduzir o sujeito à interiorização dos signos, de forma que “todas as funções superiores originam-se das relações reais entre indivíduos” (VYGOTSKY, 1998, p. 58), contrapondo as perspectivas inatistas existentes.

Bronckart (2006, p. 33) afirma que Vygotsky tratou “de demonstrar como o social se transforma em ideacional e como, subsequentemente, o ideacional interage com o corporal”. Uma vez que o indivíduo carece das relações sociais, é na interação com o outro e com os elementos socialmente estabelecidos que a noção de desenvolvimento surge como emergência da interiorização do processamento cognitivo. Vygotsky (1998), portanto, concebe como uma nova abordagem o fato de que o sujeito, a partir de um aprendizado, enquanto subsídio, desencadeia processos desenvolvimentais. Logo, os dois conceitos são definidos como dinâmicos, influenciando um ao outro, o que nega, assim, a prevalência de um ou do outro e reforça

a relação de interdependência entre eles. A proximidade inicial entre os conceitos apresentados reforça a perspectiva da Teoria da Complexidade, a qual aponta para a ideia de que é na relação do todo com as partes e vice-versa que se pode observar a totalidade de um determinado objeto de estudo, neste trabalho. Diante dessa abordagem, evidencia-se o fato de que as correntes teóricas antecessoras ao posicionamento de Vygotsky elencavam processos de forma isolada, mas, ao passo que os conceitos de aprendizagem e desenvolvimento são assumidos como interdependentes, Vygotsky (1998) não admite a perspectiva reducionista anterior e adota um olhar complexo, segundo a tomada de posição deste estudo. Nesse caso, parece reafirmar o humano e a linguagem como SACs, constituídos e constituidores de processos não lineares, interdependentes, dinâmicos, caóticos, influenciados, influenciáveis, etc

Encarando sua inquietação, Vygotsky (1998) opõe-se à metodologia investigativa, vigente em sua época, em que os testes aplicados apresentavam apenas a condição atual da criança e não permitiam intervenção a respeito dos conhecimentos que ela possuía, a fim de que, para além desses, outros parâmetros também fossem considerados. Dessa maneira, Vygotsky (1998) infere que o foco não deveria estar nos resultados enquanto produto, mas que era necessário debruçar-se sobre o processo como um todo: compreender de que maneira a criança atingia os resultados coletados, qual era a condição atual de seus conhecimentos, de que forma com o auxílio de um par mais experiente poderia manifestar a continuidade de processos ainda em maturação e que influências sua trajetória apresentava na construção desse processo (SILVA, 2009).

Neste horizonte, traçam-se, aqui, algumas inferências de Vygotsky (1998) acerca de tal concepção, de maneira a ilustrar o fato de que, mesmo não tendo por base teórica a Teoria da Complexidade, suas proposições se enquadram na complexidade inerente a todos os sistemas (cognitivo, motor, etc.) que compõem o desenvolvimento do humano. Ao conceber a interdependência do meio (contexto interacional) em relação à atividade humana, Vygotsky (1998, p. 62) afirma:

A abordagem dialética, admitindo a influência da natureza sobre o homem, afirma que o homem, por sua vez, age sobre a natureza e cria, através das mudanças nela provocadas, novas condições naturais para sua existência.

A partir disso, entende-se que os sujeitos são produtos e produtores de suas interações com o meio em que se situam, o que evidencia que sistemas se acoplam a

subsistemas e assim por diante. Por exemplo, gerações anteriores às gerações atuais serviram de modelo para suas sucessoras, quanto a alguns parâmetros socialmente construídos, mas, por outro lado, as novas gerações trouxeram à tona a emergência de novas concepções (éticas, estéticas, culturais, educacionais, linguísticas, etc.) frente aos padrões comportamentais que essas apresentavam. Ademais, nessa ótica, tantos outros elementos poderiam ser elencados na relação dialética entre o meio e o humano. Assim, sob a ótica da complexidade, poder-se-ia afirmar que o posicionamento de Vygotsky (1998) situa características pertencentes aos SACs, sendo a de maior escopo a noção de processo, ou seja, fluxo desenvolvimental contínuo e inter-relacionado ao contexto e demais sujeitos.

O posicionamento inicial de Vygotsky (1998) assume a ideia de que a ciência clássica, intitulada reducionista, vinha mutilando, isolando e refutando todos os aspectos/parâmetros/elementos/subsistemas periféricos ao objeto de estudo, ou seja, padrões não atestáveis, portanto, não replicáveis. Logo, eram irrelevantes, descartados como não sendo parte constitutiva de um todo sistêmico. Ao se assumirem as postulações de Vygotsky, infere-se que seu método investigativo considera que todos os fenômenos devem ser estudados como processos em movimento e em mudança. Nesse sentido e concernente à atividade de linguagem, Vygotsky (1998) adota o conceito de interação como aquele que decorre da atividade, por meio de ações entre o humano e o ambiente/contexto com a utilização de signos, de modo que os sistemas de signos (linguagem, escrita, numérico), tal qual o sistema de instrumentos, são estabelecidos pela sociedade e apresentam transformações no nível do desenvolvimento coletivo e individual.

É relevante instaurar a noção de Vygotsky (1998) de que a transformação de um processo interpessoal num processo intrapessoal está interpretada como “o resultado de uma longa série de eventos ocorridos ao longo do desenvolvimento” (p. 58). Na ótica da proposta teórica deste estudo, o que se ressalta é a transformação da constituição do pensamento consciente das concepções externas para internas e, não necessariamente, o período de tempo imprescindível para essa transformação. Não haveria reducionismo cartesiano maior, frente à base epistemológica desta pesquisa, se aqui fosse defendida a ideia de que diferentes sujeitos e internalizações do pensamento deveriam acontecer no mesmo nível e tempo, para todos, de maneira linear.

Na continuidade desse raciocínio, Vygotsky (1998) considera que a internalização de formas culturais de comportamento abarca a reconstrução psíquica do

sujeito na relação com os signos que o cercam. De acordo com Vygotsky (1998), devem-se analisar processos e não objetos, constatação essa que reafirma a motivação do estudo aqui proposto, pois a observação do humano e da linguagem, em perspectiva sistêmica desenvolvimental, faz com que o estágio inicial do um sujeito, sua atividade de linguagem na relação com os signos, possa desvelar padrões que, se derivados de uma observação reducionista e apenas descritiva, não possibilitaria a interpretação das relações dinâmico-causais reais subjacentes aos fenômenos da construção do pensamento humano.

Essa análise não deve estar focada apenas em uniformidades, como outrora mencionado, pois seria leviano negar a presença de padrões e comportamentos periféricos complexos às uniformidades de um determinado objeto de estudo. Adotando esse posicionamento, Vygotsky (1998, p.73) afirma que alguns psicólogos estariam equivocados ao assumirem os parâmetros fossilizados e diz que “esses psicólogos não estão interessados nas reações complexas como um processo de desenvolvimento”. No que tange à afirmação de Vygotsky, este estudo assume a necessidade de que as bifurcações comportamentais, ou seja, variações, sejam arroladas com vistas a compreender o todo processual de um fenômeno. Portanto, o desenvolvimento humano é um processo dialético complexo, que se caracteriza pela periodicidade, desigualdade no desenvolvimento de funções diversas, ou seja, é uma metamorfose contínua que engloba processos internos e externos e processos adaptativos em práticas situadas (SILVA, 2009).

De acordo com Gleick (1989), esses componentes/comportamentos deveriam ter sido assumidos enquanto fractais de um todo, pois a relação das partes com o todo, em ótica fractalizada, tem muito a revelar da constituição de um objeto de estudo. O interesse, dessa maneira, estava centrado no estado do objeto, e não, em seu processo de constituição, trajetórias e níveis possíveis de aprimoramento das condições atuais/reais/iniciais. Para Morin (2008), como já mencionado, negar a complexidade do humano é negar o humano em si. Em síntese, a citação apresentada refere-se a inter-relação, interdependência, imprevisibilidade e a processos de emergência que o humano manifesta quando observado sob uma ótica complexa que o situa em um contexto interacional específico. Dessa maneira, em seu método investigativo, Vygotsky (1998, p. 62) determina que “a aceitação dessa proposição significa termos de encontrar uma nova metodologia para a experimentação psicológica”. Conclui-se, então, que todo método de estímulo-resposta apresenta inadequidades.

Dessa forma, passam-se a observar processos e não mais objetos, em uma tomada de consciência de que, outrora, os objetos foram tomados como estáveis e fixos e de que seria adequado observar a dinâmica dos principais pontos da história dos processos. Vygotsky (1998, p. 64) assume, então, o método de desenvolvimento experimental, inferindo que, se a análise do objeto for substituída pela do processo, “então a tarefa básica da pesquisa obviamente se torna uma reconstrução de cada estágio no desenvolvimento do processo: deve-se fazer com que o processo retorne aos seus estágios iniciais”. Diante desse método, Vygotsky reforça que a explicação do processo em si não basta, sendo necessária a descrição do processo, visando a revelar suas relações dinâmico-causais reais subjacentes ao fenômeno (SILVA, 2009).

Consequentemente, Vygotsky (1998), diante de suas postulações, visa a mostrar a essência dos fenômenos psicológicos em vez de suas características perceptíveis. Como o que interessa é o processo, mais do que o estado do sujeito, Vygotsky (1998, p. 65) manifesta a necessidade de considerar os pontos de vista fenotípico (descritivo) e genotípico (explicativo) quanto ao objeto de estudo (o sujeito), a fim de desvelar sua gênese e suas bases dinâmico-causais. O autor ainda afirma que

dois processos fenotipicamente idênticos ou similares podem ser radicalmente diferentes em seus aspectos dinâmico-causais e, vice-versa, dois processos muito próximos quanto à sua grandeza dinâmico-causal podem ser muito diferentes fenotipicamente (VYGOTSKY, 1998, p.65).

Nesse entendimento, está evidente que não é suficiente compreender o objeto de estudo atrelando-o a uma constituição padronizada, portanto, reduzida, mas que é necessário investigar/acompanhar o processo e revelar quaisquer possíveis trajetórias diferentes e/ou idênticas, as quais conduziram dois processos à emergência de um padrão ou de padrões diferentes. Pertinente a esse construto de Vygotsky (1998) é sua compreensão acerca da análise, a qual descreve, afirmando que

a análise objetiva inclui uma explicação científica tanto das manifestações externas quanto do processo em estudo. A análise não se limita a uma perspectiva do desenvolvimento. Ela não rejeita a explicação das idiosincrasias fenotípicas correntes, ao contrário, sudordina-as à descoberta de sua origem geral.

Em resumo, o novo método posiciona-se como faz a Teoria da Complexidade, a qual concebe o sujeito enquanto SAC, em caráter de interdependência com elementos

físicos e psíquicos, com os quais se relaciona. Assim, o processo desenvolvimental do sujeito não pode ser isolado, bem como não se pode omitir a ideia de que diversos outros SACs influenciam e são influenciados continuamente. Portanto, observar e analisar por meio de explicações suscita verificar o ponto de metamorfose/bifurcação/mudança/transição ao qual, em geral, todos os SACs são submetidos, apresentando padrões de emergência ou retroalimentando-se de sua própria constituição, com o intuito de não deixar que o sistema “morra”. Comungando com a noção de mudança, tanto em caráter fenotípico quanto genotípico, Vygotsky (1998, p. 68) diz que “estudar alguma coisa historicamente significa estudá-la no processo de mudança: esse é o requisito básico do método dialético”. Essa concepção de base teórico-epistemológica da Teoria da Complexidade parece assegurar o pensamento complexo (MORIN, 2008), como posicionamento ontológico e praxiológico do humano. Primeiro, como anteriormente mencionado, o conceito de processo dialético complexo, como aquele que, ao ser observado, pode revelar as condições iniciais de um SAC, evidenciando suas especificidades (histórico sociocultural, experiências, falhas, progressos, fragilidades, potencialidades, etc.). O que Vygotsky (1998) define como metamorfose qualitativa pode ser definido como emergência (SILVA, 2009). De acordo com Morin (1997, p.105), os comportamentos, padrões e características que emergem dos processos denominam um processo de emergência, a qual o autor define, afirmando que

A emergência é um produto da organização que, embora inseparável do sistema enquanto todo aparece não só ao nível global, mas eventualmente também ao nível dos componentes. Assim, as qualidades inerentes às partes no seio dum dado sistema estão ausentes ou virtuais quando estas partes se encontram isoladas; só podem ser adquiridas ou desenvolver-se pelo e no todo [...].

Na mesma perspectiva, Larsen-Freeman e Cameron (2008, p.59), referindo a essas transformações/mudanças na estrutura psíquica ou comportamental, detalham:

Às vezes a auto-organização conduz a um novo fenômeno a uma escala ou nível diferente, um processo chamado ‘emergência’. O que emerge como resultado de uma fase de mudança é algo diferente da fase anterior: um todo é mais do que a soma de suas partes e não pode

ser explicado de forma reduzida por meio da atividade das partes que o compõem.<sup>19</sup>

Sustentando tal posicionamento, Larsen-Freeman e Cameron (2008, p.199) e Leffa (2009) ressaltam a visão de subsistemas, os quais se acoplam uns aos outros e ilustram tal proposição tendo a sala de aula de línguas como exemplo.

A co-adaptação parece ser um tipo relevante de mudança no sistema dinâmico de uma sala de aula de línguas. As salas de aula de línguas estão cheias de pessoas se co-adaptando: professores com estudantes, estudantes com outros estudantes, professores e estudantes com seus contextos de aprendizagem. Padrões estabilizados de ação, incluindo a ação linguística, emergem da co-adaptação de várias escalas de tempo.<sup>20</sup>

Vygotsky (1998) refere-se, também, aos embricamentos dos fatores internos e externos e processos adaptativos que superam os impedimentos que a criança encontra no que tange tal pensamento e à mutabilidade de padrões. Larsen-Freeman e Cameron (2008, p. 8) inferem:

Nós certamente aceitamos o fato de que os humanos fazem escolhas momentâneas, a fim de empregar recursos semióticos, incluindo os linguísticos, eles precisam perceber suas metas transacionais, interpessoais, auto-expressivas etc. e as múltiplas dimensões de estados de consciência e identidade, afeto e papel social.<sup>21</sup>

Ao encontrar-se em situação de turbulência/instabilidade/desordem/caos, um SAC, neste estudo, tanto o humano quanto a linguagem, busca, a partir de suas condições internas inter-relacionadas às condições externas, atingir uma condição de certo equilíbrio, de onde, retomando o conceito mencionado, surgirá a emergência. Larsen-Freeman e Cameron (2008, p. 34) ressaltam que mesmo “sistemas em equilíbrio

---

<sup>19</sup> Do original: “*Sometimes self-organization leads to new phenomena on a different scale or level, a process called ‘emergence’. What emerges as a result of a phase shift is something different from before: a whole that is more than the sum of its parts and that cannot be explained reductively through the activity of the component parts*” (LARSEN-FREEMAN; CAMERON, 2008, p. 59).

<sup>20</sup> Do original: “*Co-adaptation appears to be a particularly relevant type of change in the dynamic systems of the language classroom. Language classrooms are full of people co-adapting teacher with students, students with each other, teacher and student with learning contexts. Stabilized patterns of action, including language action, emerge from co-adaptation on various timescales*” (LARSEN-FREEMAN; CAMERON, 2008, p.199).

<sup>21</sup> Do original: “*We certainly accept the fact that humans make choices in the moment to employ the semiotic resources, including linguistic ones, they at hand to realize their transactional, interpersonal, self-expressive, etc. goals and the multiple dimensions of self and identity, affective states, and social face* (LARSEN-FREEMAN; CAMERON, 2008, p. 8).

estão continuamente adaptando-se às mudanças contextuais e podem mudar internamente como resultado de adaptação à mudança externa”<sup>22</sup>.

Na citação apresentada, muito embora Vygotsky (1998), não mencione o conceito de aprendizagem, toma-o como inter-relacionado ao desenvolvimento. Para isso, o autor traça as seguintes assertivas, apresentando ambos os conceitos:

O aprendizado é mais do que a aquisição de capacidade para pensar; é a aquisição de muitas capacidades para pensar; é a aquisição de muitas coisas. O aprendizado não altera nossa capacidade global de focalizar a atenção; em vez disso, no entanto, desenvolve várias capacidades de focalizar a atenção sobre várias coisas (pp. 92-93).

O desenvolvimento nas crianças nunca acompanha o aprendizado escolar da mesma maneira como uma sombra acompanha o objeto que a projeta. Na realidade, existem relações dinâmicas complexas entre os processos de desenvolvimento e de aprendizagem, as quais não podem ser englobadas por uma formulação hipotética imutável (p.104).

Vygotsky (1998) torna a assumir a dinamicidade presente nos processos, na medida em que cada um deles está subentendido na aproximação aqui proposta como SACs que, devido a sua constituição, apresentam dinamicidade interacional entre seus componentes/partes/subsistemas, caracterizados por apresentarem espaços de fase/estado (LARSEN-FREEMAN; CAMERON, 2008, p.48), concepção essa que será apresentada a seguir. Tanto o processo de desenvolvimento quanto o de aprendizagem, ao se depararem com fatores dinâmicos e adaptativos, permitem que os sujeitos possam emergir ao longo do processo em relação às especificidades de suas condições iniciais existenciais. Tanto é que Vygotsky (1998, p.95) afirma que, “de fato aprendizagem e desenvolvimento estão interrelacionados desde o primeiro dia de vida da criança”. A fim de aprofundar a dimensão do aprendizado escolar enquanto desenvolvimento cognitivo (funções superiores), Vygotsky (1998) propõe o conceito de Zona de Desenvolvimento Proximal (ZDP). Porém, antes mesmo de aprofundar os conceitos de zonas, que Vygotsky (1998) apresenta em dois níveis, cabe enfatizar que a denominação zonas foi a maneira que o autor encontrou para realizar a operacionalização do conhecimento científico, ou seja, para que esse não residisse apenas em uma abordagem filosófica.

Para tanto, reforça-se a ideia de que múltiplas são as zonas/estados de fase dinâmicos, interrelacionados e em constante mutação que um SAC apresenta ao longo

---

<sup>22</sup> Do original: “*Systems in equilibrium are continually adapting to contextual changes, and may change internally as a result of adapting to external change*” (LARSEN-FREEMAN; CAMERON, 2008, p. 34).

de seu desenvolvimento, o qual inicia na concepção de sua existência e transcende sua “morte”, pois, interiorizado na memória de outrem e nas ações concretizadas no mundo físico e psíquico. Seguindo esse pensamento, no que tange às zonas, aos níveis que um SAC apresenta, Vygotsky (1998, p. 95), ao debruçar-se sobre a delimitação de dois níveis, não desconsidera a existência dinâmica e complexa presente na multiplicidade de níveis que o ser humano comporta e, para isso, apregoa que “temos que determinar pelo menos dois níveis de desenvolvimento”. Indiscutivelmente, sua assertiva dialoga e compartilha de um posicionamento que está presente nas bases que fundamentam a Teoria da Complexidade. Assim, Vygotsky (1998, pp.95-96) define tais níveis, sendo o primeiro chamado de

Nível de Desenvolvimento real, isto é, o nível de desenvolvimento das funções mentais da criança que se estabeleceram como resultado de certos ciclos de desenvolvimentos já completados.

Em síntese, trata-se daquilo que a criança consegue realizar quanto a uma tarefa apresentada, sendo o resultado disso um indicativo de sua capacidade mental. Nesse caso, se outro, que não a criança, iniciar a tarefa, dando-lhe pistas para que a conclua e essa não conseguir realizá-la, então não se entenderão suas manifestações como indicativas de sua capacidade mental (SILVA, 2009). Passa-se, a seguir, ao conceito de Zona de Desenvolvimento Real (ZDP), a qual é, assim, explicada por Vygotsky (1998, p. 98):

a Zona de Desenvolvimento Proximal define aquelas funções que ainda não amadureceram, mas que estão em processo de maturação, funções que amadurecerão, mas que estão presentemente em estado embrionário.

Frente à conceituação acerca das duas zonas/níveis, afirma-se que ambas, na realidade, são, inicialmente, estados de fases, ou seja, dão origem a outros sistemas e subsistemas complexos situados no âmago do desenvolvimento humano. Essas fases têm o humano como portador de suas manifestações, enquanto zonas mais ou menos equilibradas, dependendo de suas interações. Na verdade, não se quer diminuir a complexidade de cada zona, muito pelo contrário, quer-se afirmar que essas são, na realidade, constitutivas dos processos turbulentos pelos quais o humano vai estabelecendo sua trajetória. Esse movimento contínuo pode ser considerado um eterno movimento de aprendizagem e desenvolvimento. Desse modo, Vygotsky (1998),

assertivamente, diz que o nível de desenvolvimento real caracteriza o desenvolvimento mental retrospectivamente, enquanto a zona de desenvolvimento proximal caracteriza o desenvolvimento mental prospectivamente.

Nessa ótica, é possível conjecturar que os termos relacionados aos conceitos de retrospicção e prospecção derivam-se de um SAC, o humano. A prospecção encontra-se mais como uma possível emergência do SAC, ao passo que a retrospicção está concretizada nas experiências passadas. Ao assumir, a ZDP como um estado de fase (podendo, sem dúvida, também ser analisado enquanto sistema e/ou subsistema), está-se, de fato, sugerindo que um sistema passa por uma fase de mudança quando seu comportamento muda de repente, de forma suave ou radical, para uma nova zona ou modo de interação com os elementos que o compõem e se interconectam (LARSEN-FREEMAN e CAMERON, 2008), manifestando um panorama dos possíveis estados/níveis de fase do SAC. Concernente a esse posicionamento, antes de tomar as referidas zonas como SACs, dada sua complexidade sistêmica, o que se quer neste estudo é cunhar a ZDR e a ZDP como estados de fase turbulentos.

De acordo com Larsen-Freeman e Cameron (2008, p. 48), estados de fase são estados que se inter-relacionam, demonstrando o estado atual do sujeito, manifestando as trajetórias já percorridas ou a serem percorridas. Dessa maneira, a zona de desenvolvimento real implica considerar trajetórias já traçadas; já a zona de desenvolvimento proximal suscita considerar os possíveis trajetetos a serem percorridos no caso de uma mediação formativa, rumo aos objetivos a que visa o mediador/professor, nesse caso, em determinada situação de aprendizagem, por meio da interiorização de signos. Essas zonas ou estados de fase se metamorfoseiam, alterando seus padrões de acordo com as influências dos fatores individuais e sociais que as afetam. Por vezes, elas atingem períodos de turbulência caótica, os quais as remetem a mudanças ou períodos mais instáveis, ou seja, de não fixidez, estando à beira do caos, em estado caótico (SILVA, 2009).

Essas noções dialogam com a abordagem de Vygotsky (1998) no que se refere à compreensão de que o sujeito, tanto a respeito de seu aprendizado quanto a seu desenvolvimento, necessita de pares mais experientes para alcançar formas superiores de pensamento a respeito de determinada atividade, seja ela em caráter formal ou informal. A indissociabilidade do individual e do coletivo também é mencionada por Morin (2008, p. 126), quando ele se refere ao conceito de sociedade, afirmando que

a sociedade, por exemplo é produzida pelas interações entre os indivíduos que a constituem. A própria Sociedade, como um todo organizado e organizador, retroage para produzir os indivíduos pela educação, pela linguagem, pela escola. Assim, os indivíduos nas suas interações, produzem a Sociedade, que produz os indivíduos que a produzem. Isso faz-se num circuito espiral através da evolução histórica.

A partir da conceituação das zonas mencionadas, é possível considerar que, ao se visualizar uma zona como de desenvolvimento proximal, percebe-se, mesmo que os resultados sejam imprevisíveis, que, ainda assim, é factível estabelecer condições para a instrumentalização do sujeito, para fins de aprendizagem/desenvolvimento, de modo que esse transite de uma zona real para uma zona proximal de construção de conhecimento. Isso se dá por meio de intervenções mediativas/formativas capazes de afetar o sujeito para que esse tenha subsídios para gerar a mudança frente a seus fatores internos (psíquicos) e externos (sociais). Para Vygotsky (1998, p. 98), a zona de desenvolvimento proximal “define funções que ainda não amadureceram, mas que estão em processo de maturação, funções que amadurecerão, mas que estão presentemente em estado embrionário”.

Identificar, a partir das condições reais do sujeito, o que outrora ele internalizara é inferir acerca de sua relação interacional com outros sistemas (cognitivo, motorio-sensor, linguístico, etc.), sendo essa metodologia uma retrospectiva aos fatores de sistemas e subsistemas que já o afetaram, condicionando-o a seu estado real de desenvolvimento. Por outro lado, a zona de desenvolvimento proximal situa seu desenvolvimento mental de forma prospectiva (SILVA, 2009). A análise de suas condições permite, então, verificar o que esse já possui e o que ainda pode absorver tendo suas interações formais mediadas. Neste tocante, Vygotsky (1998, p. 98) diz que a zona de desenvolvimento proximal possibilita que possamos delinear e propiciar “o acesso não somente ao que já foi atingido através do desenvolvimento, como também àquilo que está em processo de maturação”.

Enfatiza-se que a ideia de delinear também está abarcada pela Teoria da Complexidade, pois seria um reducionismo dizer que se poderia definir um estado concreto a ser alcançado em um SAC, devido às imprevisibilidades dinâmicas que o constituem. Nessa perspectiva, entende-se, frente ao aporte teórico adotado, tanto no campo da Linguística Aplicada, como no da Teoria da Complexidade, que a interiorização dos signos, enquanto aprendizado, é possibilitadora de práticas interacionais, na medida em que o desenvolvimento é permeado pelo uso de tais

interiorizações, mesmo antes da inserção do sujeito em contexto de aprendizagem e desenvolvimento formal, neste caso, escolar/acadêmico. Neste tocante Vygotsky (1998, p.97) propõe o conceito de Zona de Desenvolvimento Proximal, o qual é definido como

a distância entre o nível de desenvolvimento real, que se costuma determinar através da solução independente de problemas, e o nível de desenvolvimento potencial, determinado através da solução de problemas sob a orientação de um adulto ou em colaboração com companheiros mais capazes.

Essa concepção evidencia a necessidade de afetar para mudar, de ser afetado para organizar, desorganizar e reorganizar as funções superiores do pensamento de forma a permitir um deslocamento. O caráter social é o responsável pela metamorfose do pensar, fazer e situar-se a partir de um novo padrão cognitivo-comportamental. Incansável seria elencar todos os elementos presentes na própria constituição de cada zona. Embora essa não seja uma pretensão, compreende-se que a composição de cada zona implica uma relação dicotômica, de interdependência, para que elas coexistam e possam ser consideradas como produtoras de emergências do agir humano.

A citação de Vygotsky (1998) corrobora que há a interdependência de elementos centrais, como o cenário, o sujeito, a atividade de linguagem, a mediação formativa, os processos externos (pré-construídos socialmente), etc., ou seja, a relação sistêmica, da dicotomia todo e suas partes (MORIN, 2008). Em se tratando do conceito de ação de linguagem, anteriormente apresentado, cabe focalizar, por fim, mais um aspecto relacionado à linguagem que diz respeito à noção dos pré-construídos e de intertexto, ambos inseridos em uma arquitetura maior.

Essas ações e atividades languageiras (BRONCKART 2006; 2008) são socialmente construídas e estabelecidas pela atividade de linguagem (gênero de discurso), sendo essa constituída por vozes múltiplas que orientam o comportamento e a psique do sujeito. Não obstante a essa corroboração da assertiva de Vygotsky (1998), reafirma-se a influência entre sistemas, neste caso, sendo os seguintes: professor (ação mediativa), aluno (mediado por fatores internos e externos à sua ação), ZDR, ZDP, aprendizado e desenvolvimento, todos interconectados, interagindo rumo à manifestação da interiorização de formas superiores do pensamento e a ação realizada de concretização da atividade proposta como emergência do SAC, o aluno.

Na idealização de que o sujeito manifeste a concretização correta da atividade, tendo esse enquanto SAC sido tensionado, por meio de turbulências, a agir, também se

compreende que, nem sempre, ele realizará a atividade com maestria. Para essa conjecturação, adota-se a noção de bifurcação que, segundo Larsen-Freeman e Cameron (2008, p. 128), pode encaminhar o sujeito para interpretações outras que não aquelas almejadas por seu mediador/professor. Larsen-Freeman e Cameron (2008, p. 128) definem o conceito de bifurcação como aquele que “ocorre sempre que um sistema pode estar em dois estados ou estágios qualitativos diferentes ao mesmo tempo”<sup>23</sup>. É necessário que essa “não realização correta/coerente” seja, sim, interpretada, também, como uma manifestação de emergência da ação de linguagem do autor/sujeito/aluno de um texto (SAC), uma vez que a trajetória, aparentemente, equivocada, ao ser analisada, em caráter educacional, pode revelar estados de fase sobre os quais o mediador/professor pode intervir, possibilitando que o aluno seja submetido a atividades que o conduzam, pelo menos, ao entendimento de seu equívoco. O que se está ressaltando é que não se trata, meramente, de uma concepção de erros e acertos, mas da construção de trajetórias que podem afetar o sujeito em maior ou menor grau, causando-lhe a manifestação de um tipo de emergência. Esta abordagem teve como objetivo realizar a aproximação dos princípios de Vygotsky (1998), de maneira a enquadrá-los na ótica complexa dos SACs (SILVA, 2009), sob a qual se debruça a Teoria da Complexidade. A seguir, apresenta-se a ferramenta de mineração de dados Sobek e suas funcionalidades.

## 2.5 A ferramenta de mineração de dados Sobek

Neste subcapítulo, inicialmente, apresenta-se a noção de mineração de dados e/ou textos e, em seguida, a ferramenta de mineração de textos SOBEK, utilizada nesta pesquisa. A mineração de textos é um método de extração de informações recorrentes em um texto, a partir de uma base de dados não estruturada ou, até mesmo, semi-estruturada, conforme postulado por Feldman e Sanger (2006). Esse procedimento situa-se em um campo multidisciplinar, tendo como referência áreas do conhecimento tais como a Informática, a Estatística, a Ciência Cognitiva e a Linguística Aplicada.

---

<sup>23</sup> Do original: “...occur whenever the system can be in two qualitatively different states or stages at the same time”. (LARSEN-FREEMAN e CAMERON, 2008, p.128).

Uma vez que a linguagem é complexa e dinâmica (LARSEN-FREEMAN e CAMERON, 2008) entende-se que a ação minerativa extrai dos textos as regularidades manifestadas pelos sujeitos em suas produções textuais, visando à identificação de padrões ou comportamentos linguísticos, baseadas em análises estatísticas, enquanto apoio o caráter semântico e sintático, em geral. Para Tan (1999), por exemplo, a mineração de dados pode extrair padrões considerados não triviais, sendo esses entendidos como uma expansão do fazer inerente à técnica. Por meio desse procedimento, pode-se obter um panorama da atividade de linguagem dos sujeitos de maneira a compreender o que fora ou não mencionado. Essa técnica enquadra-se na área de *Data Mining*, podendo também ser nominada como mineração de dados textuais (KLEMANN, REATEGUI, RAPKIEWICZ, 2011). Em geral, a aplicação de uma ferramenta de mineração tem por objetivo não apenas extrair dados de um texto, mas o entendimento das razões pelas quais alguns dados/termos foram utilizados, visando a atender a um objetivo discursivo entre emissor e receptor.

Ao apontar a maior relevância/recorrência de determinados dados, uma ferramenta de mineração textual propicia possíveis intervenções a partir do panorama oferecido, de modo que, por exemplo, em contexto de ensino, um professor pode focalizar pontos a serem desenvolvidos por seus alunos, os quais se fizeram presentes nos dados dos textos minerados. Referente a isso, uma ferramenta de mineração de textos pode evidenciar potencialidades, assim como fragilidades, na produção textual de determinado sujeito (KLEMANN, LORENZATTI, REATEGUI, 2009). Nessa perspectiva, alguns elementos são importantes na constituição de uma ferramenta de mineração textual/dados, como estes: a) disponibilidade *on-line*; b) contagem de termos de um texto; c) apresentação de todos os termos; d) apresentação de termos relevantes; e) frequência dos termos; f) relacionamentos dos termos; e g) visualização gráfica dos termos e relações estabelecidas.

Sobek é uma ferramenta capaz de extrair termos relevantes em documentos e encontrar as relações entre esses, a partir do processo de mineração de dados de um texto. Assim, o usuário/professor é capaz de identificar conceitos relevantes a partir da análise da frequência dos termos/dados. Essa ferramenta, segundo Schenker (2003), identifica os termos/conceitos recorrentes em um texto e, dessa forma gera um grafo automatizado. Segundo Feldman e Sanger (2006), quatro são os tipos de características que demarcam os dados de um texto: caracteres, palavras, termos e conceitos. Os caracteres são componentes individuais das palavras, ou seja, letras, números e

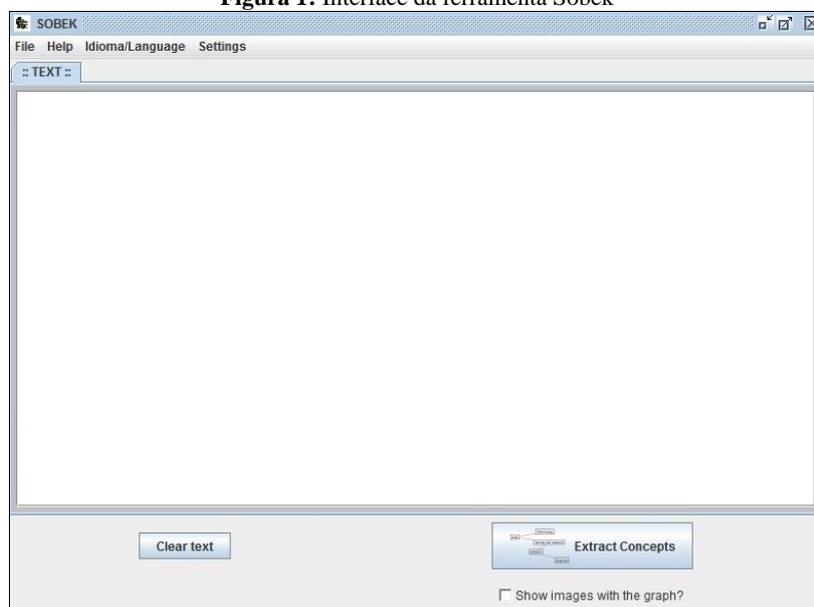
símbolos. As palavras de um texto, em geral, quando sem nenhum tipo de pós ou pré-processamento podem entendidas como forma mais básica de extração de conhecimento acerca do texto. Em geral, a redução de espaço de resposta da ferramenta é desconsiderar palavras compreendidas como pertencentes a um conjunto de “*stop words*” (grupo de palavras comuns que não acrescentam informação semântica (artigo, proposições, numerais). Os termos, por sua vez, constituem-se por uma única palavra ou conjunto, em que expressões compostas podem ser apontadas. Por fim, os conceitos são elementos gerados para um documento através de análise manual, estatística, ou com base em regras, ou abordagem híbrida (FELDMAN e SANGER, 2006).

Nessa perspectiva, de acordo com Klemann, Lorenzatti, Reategui (2009), Sobek é compreendida como uma ferramenta de apoio pedagógico quanto à escrita de textos em atividade de produção textual. Por meio do estabelecimento de um grafo constituído pelos termos recorrentes do texto-referência, o sujeito visualiza uma rede de conceitos que pode auxiliá-lo em seu processo de aprendizagem (TORRANCE e GALBRAITH, 2006), no que se refere à estruturação de um texto ou estruturação de seu discurso em uma interação com outros interagentes, acredita-se que o grafo possa auxiliá-los nessa empreitada.

A definição de grafo está compreendida como construções matemáticas importantes e eficazes para modelar estruturas de informação, com a finalidade de indicar a existência de conexões entre os dados. A função de um grafo, portanto, é a apresentar os conceitos considerados mais recorrentes, de acordo com a frequência de ocorrência deles. Entre os conceitos há, em geral, conexões, que representam relações existentes no texto. A utilização de um grafo simboliza a possibilidade de reorganizar cognitivamente o que foi lido, transpondo, de maneira clara, as noções manifestadas no texto produzido, intitulado aqui de texto-tarefa. Assim, o interagente vai percebendo a dinamicidade interna do texto por meio das relações esboçadas no grafo com os termos minerados. Blanchard e Farstrup (2011) referem-se à Sobek como uma ferramenta tecnológica de uso mais genérico, uma vez que pode ser utilizada em muitas áreas do conhecimento, não tendo como pretensão uma aplicação restrita apenas a uma área de atuação profissional. A ferramenta Sobek, desenvolvida por uma equipe multidisciplinar da Universidade Federal do Rio Grande do Sul, intitulada como Grupo de Pesquisa em Tecnologia Aplicada à Educação (GTech-Edu), sob coordenação do Prof<sup>o</sup>. Dr. Eliseo Berni Reategui, está disponibilizada na Internet, permitindo acesso gratuito a usuários que por ela se interessem. Ela é executável em diferentes sistemas operacionais (Linux,

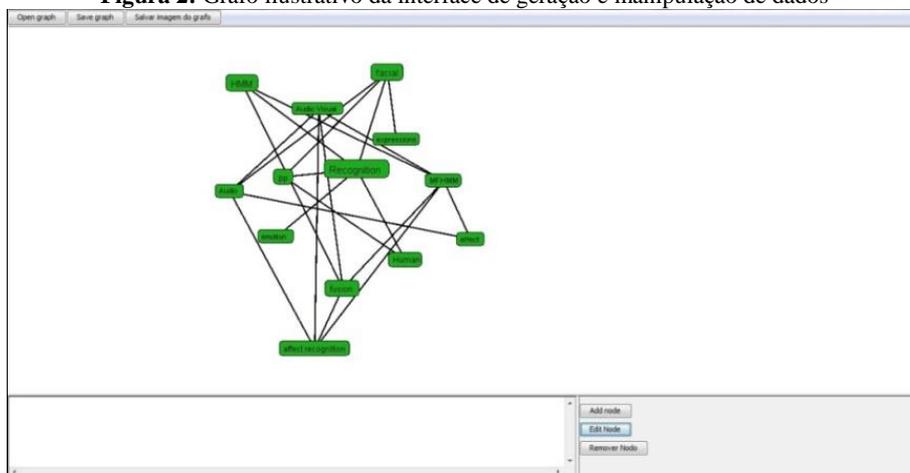
Windows ou Mac OS) e tem sido desenvolvida e aprimorada para uso em dispositivos móveis. Essa ferramenta pode ser acessada pelo endereço eletrônico <<http://sobek.ufrgs.br/>>, no qual se pode obter informações acerca de sua origem e funcionamento. Sua interface (Figura 1) é intuitiva e possibilita algumas ações sobre o que é minerado, com vistas à produção de um grafo de dados. A ferramenta permite que o usuário importe para o seu interior textos nos formatos txt, doc e pdf, para que sejam minerados, de modo a gerar um grafo com as relações estabelecidas entre termos (Figura 2). Quanto ao grafo, algumas ações são permitidas, como, por exemplo, adicionar um nodo, editar um nodo e remover um nodo (nodo, neste contexto, significa termo/dado localizado no grafo).

**Figura 1:** Interface da ferramenta Sobek



Fonte: Sobek

**Figura 2:** Grafo ilustrativo da interface de geração e manipulação de dados



Fonte: Sobek

Segundo Reategui, Pinho e Epstein (2011), a configuração presente na ferramenta baseia-se num parâmetro  $n$  para extrair os conceitos compostos com mais de uma palavra. Esse parâmetro foi criado visando a uma combinação da palavra atual com as palavras subsequentes  $n$ . Assim, tem-se uma grande combinação de palavras para encontrar o grupo mais frequente delas que aparece no texto. Por exemplo, considerando-se  $n = 3$ , a análise da sequência dos termos "AA BB CC DD EE FF GG HH" conduz às combinações "AA", "AA BB", "BB CC AA", "BB", "BB CC", "BB CC DD", e assim por diante. De modo a evitar sequências começando com preposição ou artigos, filtros específicos são utilizados. Depois de identificar as combinações mais frequentes de palavras, intituladas então de conceitos, o processo de mineração seleciona o conjunto principal de mais relevantes com base em sua frequência no texto.

Na sequência, o próximo passo é calcular a semelhança entre os conceitos. Considerando, por exemplo, dois conceitos  $x = \text{"AA DD BB"}$  e  $z = \text{"BB CC DD EE FF AA"}$ , o coeficiente de semelhança entre eles é calculado com o produto escalar também usado no *Model Space Vector*. O coeficiente de similaridade, representado por  $S$ , calcula a quantidade de palavras que estejam presentes em ambos os conceitos representados por  $P$  e o número de palavras do conceito maior representado por  $B$ . Assim, tem-se, segundo Reategui *et alii* (2011),  $S = P / B$ . Com referência ao exemplo anterior,  $S = 0,5$ , como os conceitos têm três palavras em comum, as palavras "AA", "BB" e "DD", o conceito  $z$ , sendo o maior, tem seis termos. Depois de calcular o valor de  $S$ , o coeficiente de relevância,  $R$ , é calculado para cada conceito.

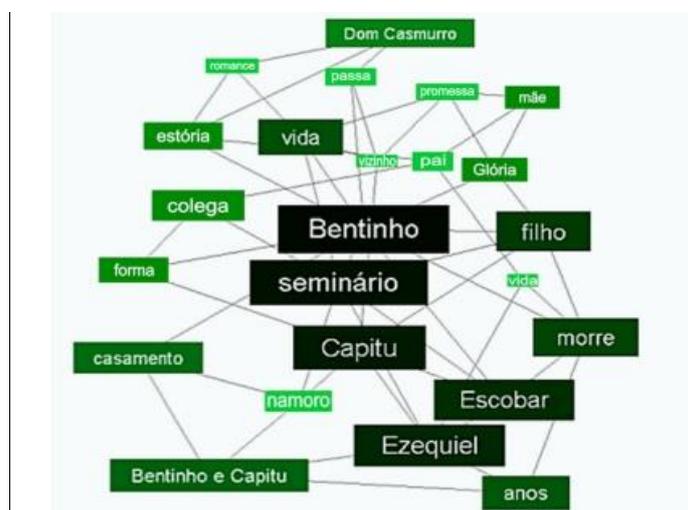
O número de palavras do conceito ( $C$ ) e a frequência absoluta ( $F$ ) são introduzidos no processo de computação. Para calcular o valor de  $R$  para cada conceito, a seguinte fórmula é utilizada:  $R = S * C + F$ . O conceito com o maior valor para  $R$  é mantido na base e, no final do processo, está incluído no gráfico. No exemplo mencionado, considera-se que o conceito tem  $x = 3$  e  $C F = 3$ , e  $Y$  tem conceito  $C = F = 6$  e  $2$ . Pode-se concluir que conceito  $z$  vai permanecer na base para ser parte do gráfico, mesmo se o seu valor  $F$  for menor do que o do conceito  $x$ . A idéia por trás do coeficiente de relevância,  $R$ , é comparar dois conceitos.

Deste modo, a ferramenta Sobek está entendida como mediadora das práticas discursivas, uma vez que pode auxiliar na produção escrita dos interagentes. Logo, ao construir um grafo automatizado, ela apresenta os termos mais relevantes de um texto e as possíveis relações, assim como termos não encadeados, mas recorrentes. Nesse tocante, cabe ressaltar que, neste estudo, a ferramenta Sobek será utilizada para fins de

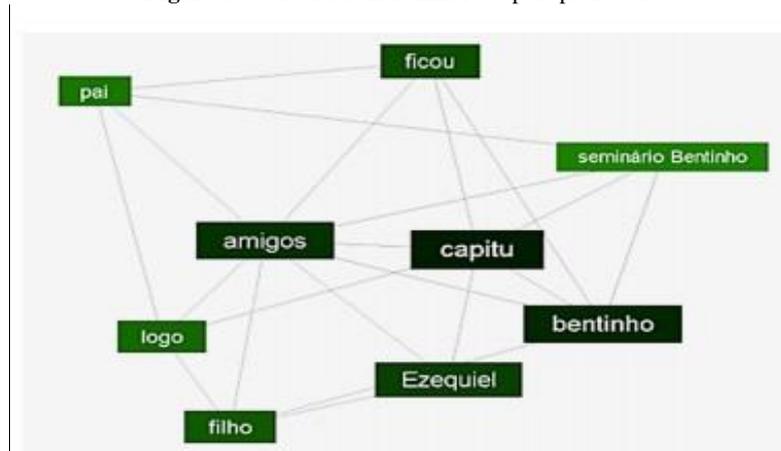
identificação de níveis de complexidade em caráter de estabilidade ou variabilidade de conceitos. Ainda na perspectiva de apoio à leitura e à produção textual, Macedo *et alii* (2011) desenvolveram um estudo quanto às estratégias pedagógicas de apoio à leitura, à escrita e ao acompanhamento, baseado na tecnologia de mineração de texto, o qual apontou para o fato de que a ferramenta Sobek se faz pertinente quanto a esses processos. Por sua vez, Azevedo *et alii* (2011) também estudaram a aplicação da ferramenta Sobek, visando à identificação da relevância de dados apresentados em mensagens postadas em fóruns de discussão, como forma de apoio à ação docente. Ambos os estudos mencionados evidenciaram a relevância da utilização da ferramenta em contextos que visem ao aprimoramento de estratégias de leitura, produção textual e acompanhamento da atividade de linguagem de sujeitos, interagindo em contexto digital de ensino e aprendizagem e/ou, na modalidade face-a-face.

Na medida em que o interesse deste estudo é focado no processo minerativo de identificação de níveis de complexidade, a seguir apresentam-se, brevemente, as inferências de Macedo *et alii* (2011) referentes ao acompanhamento de produções textuais quanto ao processo de mineração de um texto-referência, ofertado para leitura, assim como a geração de seus grafos em relação ao texto-tarefa (a produção textual). Ao comparar o grafo gerado pela ferramenta de mineração textual Sobek, a partir dos dados extraídos do texto-referência (Figura 3) em relação ao grafo do texto-tarefa (Figura 4), Macedo *et alii* (2011) afirmam que, por meio da comparação/análise “manual”, é possível inferir que o sujeito autor da produção textual: a) considerou os aspectos essenciais do texto-referência; b) apresentou um texto coerente e c) pareceu demonstrar domínio do conteúdo lido.

**Figura 3:** Texto-referência minerado pelo professor



Fonte: Macedo et alii (2011)

**Figura 4:** Texto-referência minerado pelo professor

Fonte: Macedo et alii (2011)

O estudo apresentado por Macedo *et alii* (2011) demonstra que a ferramenta favorece o andaimento (VYGOSTKY, 1998) por meio da abordagem tecnológica e pelas possíveis intervenções que o professor pode realizar diante do processo de mineração textual. No próximo capítulo, descreve-se a metodologia de pesquisa.

### 3 METODOLOGIA

Esta é uma pesquisa quali-quantitativa (CRESSWELL, 2007), que tem por base as concepções da Teoria da Complexidade (MORIN, 2000, 2008; LARSEN-FREEMAN, 1997; LARSEN-FREEMAN e CAMERON, 2008), inserido nos modos de pensar o recorte gerado pela ferramenta de mineração de dados Sobek, sob uma ótica interdisciplinar, compreendendo-se que esse recurso digital pode promover uma interpretação mais complexa acerca da ação de linguagem (BRONCKART, 2006, 2008) de sujeitos engajados em atividades de produção textual. A abordagem quanti-qualitativa, neste estudo, está implicada sob duas perspectivas: a) qualitativa – por se propor a postular, com base teórica, sobre como a análise de um grafo pode ser concebida, enquanto recorte tecnológico, que pode permitir a identificação da ação de linguagem atrelada a possíveis níveis de complexidade, que se configuram como uma estratégia metodológica de aprofundamento de um panorama inicial apresentado pela ferramenta; b) quantitativa – por sugerir que a utilização da ferramenta permite que o usuário visualize, não índices estatísticos, como propostos na análise de dados para fins de ilustração do estudo, mas conceitos que podem remeter ao entendimento de que o sujeito/aluno (ação de linguagem) manteve-se mais estável ou variável (LARSEN-FREEMAN e CAMERON, 2008) às condições iniciais de produção textual, ou seja, ao conteúdo de um texto-referência ofertado em uma prática educacional.

Em síntese, a finalidade da dimensão qualitativa é a de instaurar um posicionamento unicista entre os campos do conhecimento implicados neste estudo, ao passo que a dimensão quantitativa tem por objetivo afirmar a potência de um recurso digital, ao tornar observável a complexidade da ação de linguagem, quando materializada na forma de textos-empíricos socialmente validados em práticas coletivas, em contexto formal de ensino superior. Desse modo, este estudo parte da inferência de Larsen-Freeman e Cameron (2008, p.193), de que arrolar a ação de linguagem (discurso), como Sistemas Adaptativos Complexos (SAC) é “o primeiro passo, envolvendo a identificação de possíveis sistemas e suas dinâmicas, os agentes/elementos nos sistemas e os vários tipos de relações entre eles.”<sup>24</sup> Frente a isso, como mencionado ao longo desta pesquisa, a observação da ação de linguagem está

---

<sup>24</sup> Do original: “...is the first step, involving identification of possible systems and their dynamics, the agent/elements in the systems and the various types of relation among them”. (LARSEN-FREEMAN e CAMERON, 2008, p.193)

delimitada pelo recorte tecnologicamente observável, propiciado pela Ferramenta Sobek, uma vez que, ao extrair conceitos recorrentes, sendo este método também constituinte deste estudo, gera um grafo composto por conceitos (panorama do todo composicional do texto), o qual pode sugerir padrões comuns a essa atividade humana que é a de apropriação linguística como desenvolvimento humano. Portanto, a definição de SACs é elencada para compreender a emergência de padrões/níveis de complexidade, descrever as implicações do conteúdo propiciado pela ferramenta, como parte constitutiva dos textos (em nível de signos linguísticos), e apontar, quantitativamente, a composição dos grafos, associando esses índices aos índices do conteúdo do grafo dos textos-referência.

Cabe ressaltar que o tratamento de dados em caráter quantitativo não visa à concepção de que essa abordagem se torne replicável, dada a dinamicidade da língua/linguagem em uso, mas é adotado como estratégia de verificar tendências de estabilidade (possível conservação da forma e conteúdo) e variabilidade (mudança na forma e/ou conteúdo composicional em relação ao texto-referência). Relativamente a essa tomada de posição, entende-se, também, que, posto o caráter inaugural desta pesquisa, não se está operacionalizando a condução do estudo com modelos prescritivos, mas sim, a partir de concepções basilares que permitem e norteiam a construção de um conhecimento empírico acerca do tema em questão. Com isso, compreende-se que a pesquisa pode demandar ajustes na programação da ferramenta, a partir da constatação de seu desempenho referente a um processo de mineração que visa a comparar dados derivados de dois textos como forma de identificar e interpretar a complexidade da ação de linguagem. Frente a esse aspecto, visa-se à ampliação da concepção de compreensão acerca dos resultados da técnica de mineração, sob uma ótica complexa e ciente de que tal complexidade pode remeter os resultados do processo à identificação de tendências de desenvolvimento humano associadas, neste estudo, às zonas de desenvolvimento propostas por Vygotsky (1998).

De acordo com Larsen-Freeman e Cameron (2008), a explicação das dinâmicas que constituem um SAC (neste estudo, a ação de linguagem) é a tarefa maior do pesquisador e, para tanto, a metodologia deste estudo está assim dividida: a) Estudo preliminar e b) Estudo final. No estudo preliminar, executa-se a tentativa de examinar, por meio de análise, o modo como Sobek pode auxiliar na identificação da complexidade entre os dados de dois textos (referência e empírico), seguida da verificação da necessidade de implementação de novas funcionalidades na ferramenta

para este estudo (nova versão). Em relação ao estudo preliminar, optou-se por mantê-lo como parte do capítulo, uma vez que a ferramenta se configura como método de análise. Então, a exploração e a possível inserção de novas funcionalidades em Sobek estão vistas como partes constitutivas deste estudo, em caráter metodológico, sendo um primeiro desdobramento do fazer científico desta tese. Essa operacionalização tem por finalidade organizar, de modo objetivo, a apresentação e a análise do objeto de estudo desta pesquisa, entendendo, contudo, que o estudo preliminar, enquanto exercício inicial, está implicado como linha mestra, que norteia a realização da análise de dados de modo mais aprofundado.

O estudo final, por sua vez, objetiva a aplicação da mineração de dados com vistas à identificação de níveis de complexidade entre texto-referência e texto empírico. Portanto, nessa etapa, estão elencados os seguintes objetivos: i) analisar a utilização da ferramenta Sobek acerca do nível de complexidade propiciado pelo recorte que ela oferece na forma de grafos; ii) postular possíveis níveis de estabilidade e variabilidade do conteúdo do texto-empírico, considerando sua dinamicidade composicional em relação ao texto-referência de uma atividade de produção textual; iii) propor, a partir das possibilidades da ferramenta, formas de gerar tipos de grafos automatizados e distintos acerca de inter-relações entre dois textos; iv) discutir o processo de mineração de dados, os níveis de complexidade identificados, a fim de evidenciar o quanto eles sugerem um padrão de desenvolvimento humano. São apresentadas as especificidades de cada etapa – estudo preliminar e estudo final –, de modo a contextualizar sua condução. No próximo subcapítulo, descreve-se o estudo preliminar.

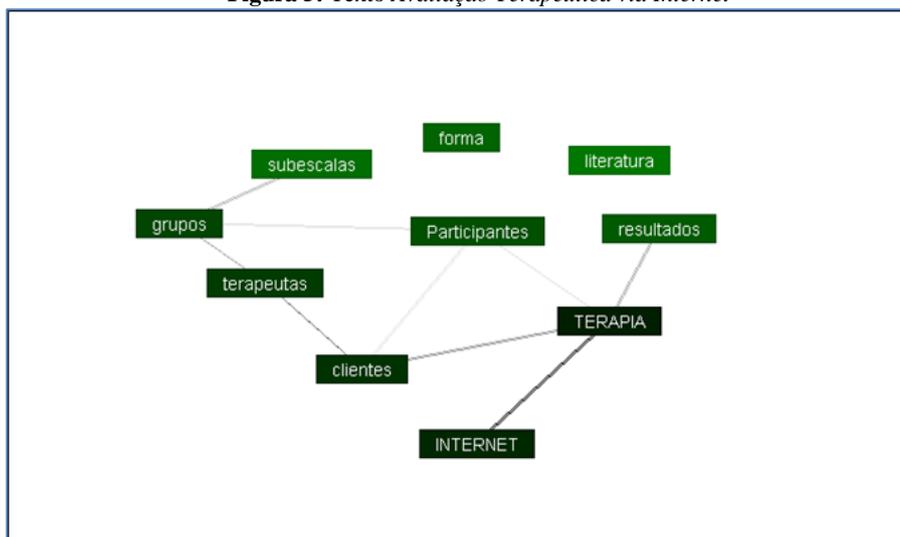
### **3.1 Estudo Preliminar**

Levando em consideração a concepção teórico-epistemológica apresentada e norteadora deste estudo, realizou-se um estudo preliminar com dados empíricos quanto ao processo de mineração, com vistas a contemplar, parcialmente, os objetivos específicos que se quer desenvolver. Para isso, utilizaram-se dados do *corpus* de pesquisa de mestrado do autor desta tese, com o objetivo de verificar se o recorte feito pela ferramenta Sobek era capaz de evidenciar padrões de estabilidade e variabilidade, conforme apregoam Larsen-Freeman e Cameron (2008) quanto à complexidade entre o

conteúdo de textos distintos. Embora não tenha sido realizada uma investigação com o mesmo tipo de texto, como na etapa do estudo final (texto-referência e texto-empírico, ambos de gênero textual da esfera acadêmica) os textos utilizados foram, de todo modo, elegíveis, como exercício inicial, para a identificação que se pretendia realizar. Portanto, o estudo preliminar envolveu 3 textos, sendo 2 textos-referência (artigos científicos), no formato dos dados oficiais coletados para tese e 1 texto-discurso, coletado de um *Chat*, em que 5 sujeitos, após a leitura de dois textos-referência, deviam interagir, de maneira a discutir o conteúdo dos textos lidos. O primeiro texto intitulava-se *Terapia pela Internet* e o segundo, *Avaliação Terapêutica via Internet*, ambos pertencentes à esfera acadêmica. Desse modo, o estudo analisou as interações discursivas oriundas de um *Chat* (2 horas de interação) de uma disciplina de graduação, do curso de Psicologia, de uma instituição de ensino superior, da rede privada, situada no Rio Grande do Sul.

Após os alunos terem lido, previamente, dois textos-referência, disponibilizados pelo professor da disciplina, na plataforma Moodle, acerca do tema *Terapia em contexto virtual*, esses interagiram em um *Chat*, tendo seu texto-discurso registrado automaticamente pela plataforma educacional, o qual, posteriormente, foi compilado e serviu de conteúdo a ser comparado aos outros dois textos. A estratégia metodológica, portanto, visava à verificação da presença de características de estabilidade ou variabilidade de conceitos. Primeiramente, os textos-referência foram submetidos à extração de dados, gerando um grafo para cada texto, e, em seguida, o texto-discurso gerou um terceiro grafo. Ressalta-se, contudo, que isso envolveu um processo manual de identificação de padrões, em que, após a geração de grafos, cada um teve seu conteúdo observado de forma isolada. Essa observação não ocorreu de forma totalmente automatizada, uma vez que os dados de cada grafo, quanto à existência de estabilidade ou variabilidade, tiveram de ser analisados pelo pesquisador, e não, pela ferramenta Sobek. Na Figura 5, tem-se o grafo extraído a partir do intitulado *Avaliação Terapêutica via Internet*.

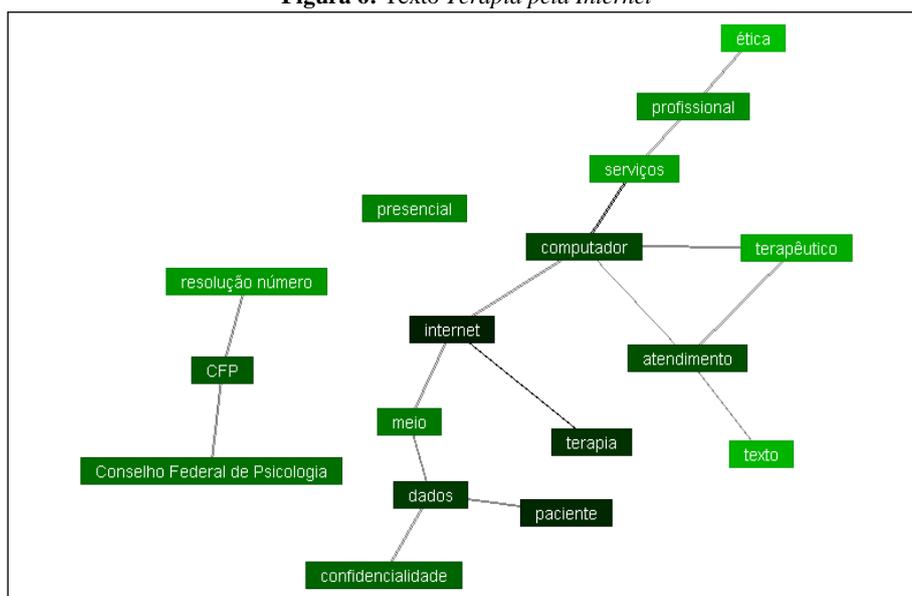
**Figura 5:** Texto *Avaliação Terapêutica via Internet*



**Fonte:** Autor

No grafo da Figura 5, é possível perceber que alguns conceitos (signos) como, por exemplo, terapia e Internet são termos mais relevantes e, portanto, recebem destaques de localização, cor e tamanho, tais como nos respectivos exemplos: grupos, clientes, terapeutas, dentre outros. De acordo com Hessler e Reategui (2010), Sobek possibilita o ato de questionar o que de fato é lido pelo autor do texto, como será discutido no grafo da Figura 7, a seguir apresentado. Já na Figura 6, tem-se o grafo extraído a partir do texto intitulado *Terapia pela Internet*.

**Figura 6:** Texto *Terapia pela Internet*



**Fonte:** Autor

No grafo da Figura 6, assim como no anterior, é possível perceber que alguns termos, como, por exemplo, Internet, terapia e paciente são mais relevantes e, portanto, recebem destaques de localização, cor e tamanho, tais como nos respectivos exemplos: computador, atendimento, dados, dentre outros. Os grafos apresentados visam a estabelecer um panorama dos termos mais frequentes nos textos minerados. Então, a fim de ilustrar esta abordagem, de maneira a aprofundar sua aplicação, realizou-se a mineração das práticas discursivas dos sujeitos em relação ao texto-empírico oriundo do *chat*. Na Figura 7, são apresentados os dados minerados do texto-discurso (texto-tarefa), o qual, posteriormente, foi comparado de forma manual aos grafos das Figuras 5 e 6, de maneira a, posteriormente, identificar a complexidade da ação de linguagem quanto a conceitos que podem ser apontados como estáveis ou variáveis ao conteúdo dos textos-referência.

A partir da análise dos textos minerados e da comparação de seus grafos (Figuras 5 e 6) em relação ao grafo do conteúdo do *Chat* (Figura 7), notou-se a ocorrência de um processo interacional que demonstrou a emergência dinâmica da ação de linguagem dos sujeitos (BRONCKART, 2006; 2008), em referência a padrões de estabilidade e variabilidade do conteúdo do texto-empírico, no que diz respeito ao conteúdo utilizado na atividade de linguagem, enquanto prática interacional. Frente a isso, pode-se dizer que, de acordo com a prerrogativa proposta pela Teoria da Complexidade, percebeu-se que os sujeitos, enquanto SACs, por intermédio da ação de linguagem, organizaram-se de forma a manter conceitos (signos) presentes nos textos-referência, porém fizeram uso de domínios do conhecimento (pré-construídos) presentes em seus arcabouços linguísticos (BRONCKART, 2006, 2008), de modo a inserir conceitos (signos) que não apenas aqueles apresentados nos textos-referência.

**Figura 7:** *Texto-discurso do Chat*



Fonte: Autor

Com isso, entende-se os textos-referência (Figuras 5 e 6) como sistemas externos à ação de linguagem dos sujeitos e de força entrópica, os quais, ao gerar uma tensão aos sujeitos, fizeram com que eles tivessem de modalizar suas ações de linguagem, sendo, assim, afetados pelo conteúdo de tais textos, dinâmica que (LARSEN-FREEMAN e CAMERON, 2008), supostamente, manifestou uma transição da Zona de Desenvolvimento Real (ZDR) para a Zona de Desenvolvimento Proximal (ZDP) (VYGOSTKY, 1998). A representação desta aparente transição deve-se ao fato de que a ação de linguagem dos sujeitos passou por um estado de fase caótico, pois fora necessário ler, interpretar, interagir e co-construir, a partir de mundos distintos (objetivo, social, subjetivo), sendo esses SACs, em relação aos mundos dos autores dos textos-referências, de forma a atingir o objetivo proposto pela atividade de leitura e interação no *Chat*. A análise manual, conduzida pelo pesquisador desta tese, evidenciou que a ferramenta Sobek, ao gerar um panorama fractalizado da relação do texto-empírico aos textos-referência, acaba por complementar ação de quem observa a constituição da ação de linguagem dos sujeitos. Assim, ela é complementar ao processo humano de interpretação. Posto que este estudo preliminar apenas vise à análise da aplicabilidade da ferramenta Sobek na identificação dos componentes dos textos, não foi objetivado um caráter quantitativo aos conceitos que, porventura fossem percebidos como estáveis aos textos-referência ou que variassem em relação a eles, apenas realizando-se o seu registro.

De maneira a ilustrar a constatação realizada, quanto à presença de conceitos estáveis e/ou variáveis, alguns termos são apresentados abaixo, na Tabela 1, evidenciando que um texto-empírico pode manifestar a emergência de escolhas linguísticas, não se limitando às condições iniciais (texto-referência), ou seja, à mera repetição dos termos lidos, mas à presença de sinônimos e novos conceitos, como um padrão de mutabilidade sistêmica.

**Tabela 1:** Termos minerados

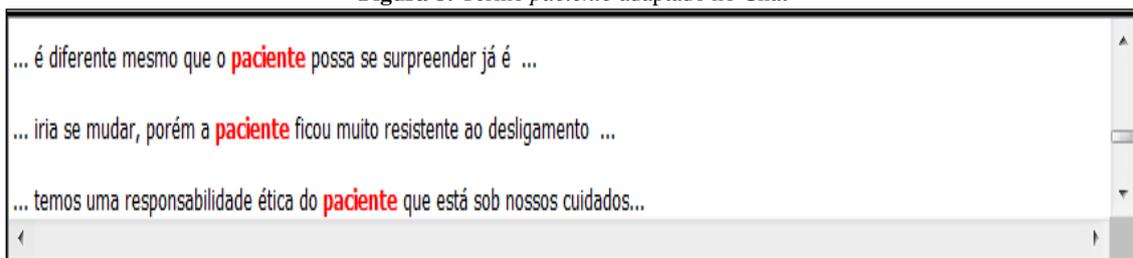
| <b>TERMOS ORIGINAIS NOS<br/>TEXTOS-REFERÊNCIA</b> | <b>TERMOS NO CHAT<br/>TEXTO- DISCURSO (TEXTO<br/>EMPÍRICO)</b> |
|---|--|
| Cliente   | Paciente   |
| Conselho Federal de Psicologia                    | Comitê   |
| Meio e Forma                                      | Virtual  |
| Internet  | A distância  |

**Fonte:** Autor

Ao término deste exercício inicial, percebeu-se que alguns termos, como os apresentados na Tabela 1, haviam sido adaptados, ou seja, tornaram-se emergentes ao serem comparados aos termos lidos quando submetidos à extração de dados. Essa adaptação corroborou o objetivo que se tinha em constatar a dinamicidade discursiva da língua/linguagem em caráter processual complexo sistêmico, ou seja, ressaltou que não se podem conceber interações discursivas como fixas ou como práticas prototípicas, como, por vezes, idealizadas por professores. Desse modo, as práticas discursivas, em relação às condições iniciais de leitura dos textos-referência, evidenciaram a presença de variações (estáveis e variáveis), dada a utilização dos termos sinônimos e novos em relação aos lidos. A recorrência dos conceitos, como nos textos-empíricos, aponta para o fato de que um SAC evidencia suas condições iniciais, por mais discretas que essas possam estar nas marcas deixadas pelo falante de uma língua (LARSEN-FREEMAN e CAMERON, 2008). Diante de tal presença, investigou-se a forma como esses termos foram utilizados por meio de um recurso oferecido pela ferramenta, o qual pode apurar ainda mais as considerações de quem os minera, como, por exemplo, um professor. Para isso, utilizou-se o recurso de visualização da recorrência de termos, enquanto auxílio na interpretação e coerência desses. Essa ação deu-se por meio do recurso que indica o número de incidência e, de forma prática, aponta o local em que se encontram, no texto minerado, por meio de uma janela que disponibiliza a estrutura sintática onde o conceito se localiza.

Assim, o usuário/professor, ao clicar em um nodo específico (conceito) disposto no grafo, faz com que Sobek ofereça a visualização de todas as vezes em que o conceito fora aplicado. Essa alternativa, complementar, possibilita que o usuário/professor avalie se a aplicação de um conceito, realmente, foi coerente e pode definir a qualidade da contribuição inserida no contrato conversacional estabelecido de maneira escrita. Visando a apresentar a utilização desse recurso, optou-se por apresentar a recorrência dos conceitos *paciente e virtual*, uma vez identificados como adaptados na análise inicial do grafo do texto-empírico. Na Figura 8, apresenta-se como nodo selecionado o termo *paciente*, adaptado no texto-empírico e considerado recorrente pela ferramenta Sobek.

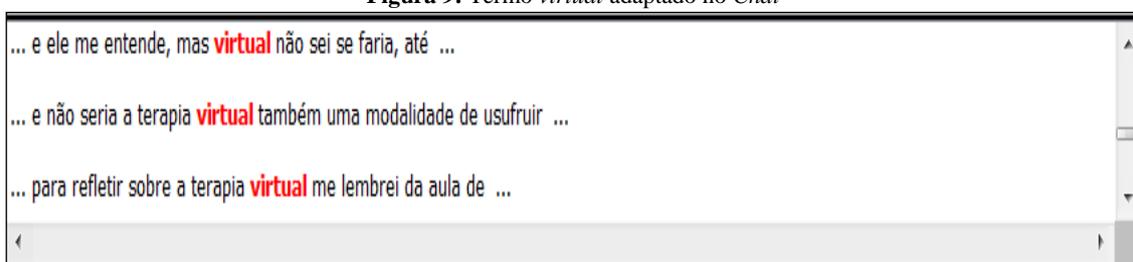
**Figura 8:** Termo *paciente* adaptado no Chat



Fonte: Autor

Na Figura 8, são exibidas três estruturas sintáticas fragmentadas com a aplicação do termo *paciente*, feitas apenas como forma de ilustrar outra possibilidade de visualização da dinamicidade da língua/linguagem em fluxo. Ao pular o conteúdo desta janela, o usuário/professor pode visualizar as demais aplicações do mesmo termo. Ao considerar que o termo *cliente*, presente nos textos-referência lidos pelos alunos, embora não possua o mesmo radical de cliente, notou-se que este fora aplicado para se referir àquele. Na figura 9, apresenta-se como nodo selecionado o termo *virtual*, que também fora considerado recorrente pela ferramenta Sobek.

**Figura 9:** Termo *virtual* adaptado no Chat



Fonte: Autor

A Figura 9 exibe três estruturas sintáticas fragmentadas com a aplicação do termo *virtual*. Diferentemente da Figura 8, em se tratando de um termo sinônimo ao conceito *meio e forma*, como apresentado na Tabela 1, é possível inferir que, se fosse substituído, haveria alteração do sentido dado, anteriormente, pelo termo virtual. Assim, se o termo fosse interpretado e entendido como resultado de um fenômeno linguístico emergente, esse poderia ser compreendido como apropriado, portanto, coerente nas recorrências exibidas por Sobek. Ressalta-se que o grau de mutabilidade, neste caso, foi superior ao da adaptação do termo *paciente*. Se, por exemplo, os termos *meio e forma* tivessem sido substituídos, pela expressão *contexto digital*, o nível de emergência linguística teria sido inferior ao que, de fato, ocorreu. O objetivo inicial deste estudo

preliminar foi o de demonstrar que a mutabilidade da linguagem pode ser observada, caso o usuário/professor esteja consciente de que a não recorrência de alguns termos presentes no texto-referência é originária do caráter fenomenológico da língua, e não, por se tratar de um uso inapropriado dos termos que emergiram do nível individual para o coletivo.

No que se refere ao caráter adaptativo e emergente identificado, entende-se que a ação de linguagem apresenta emergências, como sendo as características e os comportamentos linguísticos que emergem dos processos. Então, segundo Larsen-Freeman e Cameron (2008), o que é *mister* para este estudo é o fato de que as características manifestadas por SACs surgem “não só ao nível global, mas eventualmente também ao nível dos componentes” (MORIN, 1996, p.105). Observando-se as Figuras 5, 6 e 7, notou-se que os termos, embora em condição de estabilidade e variabilidade, apresentaram diferentes graus de adaptabilidade realizados pelos alunos. Essa análise ocorreu de forma manual, ou seja, comparando termo a termo (Figuras 5, 6 e 7), na intenção de identificar um novo padrão ou comportamento caótico e/ou emergente quanto à ação de linguagem utilizada no *Chat*. Ressalta-se que a atual versão da ferramenta Sobek não realiza este processo de forma automatizada, o que facilitaria a compreensão do usuário/professor, quanto ao processo de identificação das zonas de desenvolvimento (VYGOSTKY, 1998) fractalizadas em caráter de mutabilidade do conteúdo, para fins de mediação (BRONCKART, 2006, 2008). Em outras palavras, o que se está inferindo é que a identificação de conceitos estáveis ou variáveis revela um panorama da estrutura do texto e pode evidenciar para o usuário/professor indicativos de estratégias que esse tenha em vista, a fim de atingir os objetivos de sua disciplina, de modo que se torne possível analisar o desenvolvimento do todo (a turma), assim como em nível individual (sujeito/aluno).

Dessa maneira, ao concluir o estudo preliminar e ao compreender que a ferramenta Sobek potencializa a visualização de níveis de complexidade, na medida em que permite a identificação de conceitos (signos) que emergem em caráter de estabilidade ou variabilidade, em relação à transição e à composição de um sistema para outro (texto-referência para o texto-empírico), considerou-se relevante adaptar a versão atual da ferramenta que minera um texto de cada vez, para uma versão que suportasse a importação de dois textos e provesse maior complexidade informacional, de modo automatizado, sem ação manual de quem aplica a técnica de mineração, com vistas ao fornecimento de relações entre os seus conteúdos. Essa ação, desencadeada a partir do

aporte teórico-metodológico adotado, compreende que a função da ferramenta Sobek é a de apontar para uma característica primordial, em se tratando do conceito de SACS, que é a de tornar observável padrões de emergência do sistema, neste caso, a ação de linguagem de quem produz um texto empírico. Contudo, é essencial ressaltar que a técnica de mineração, sem dúvida, complementa a ação de quem observa (tutor/professor/pesquisador) a transição e as mudanças transcorridas da constituição de um sistema (texto-referência) para outro (texto-empírico), porém sua potência de extração dá-se no nível da estrutura, ou seja, não abarca a totalidade da composição enunciativa dos sujeitos, posto que o grafo desvela um fractal do todo textual.

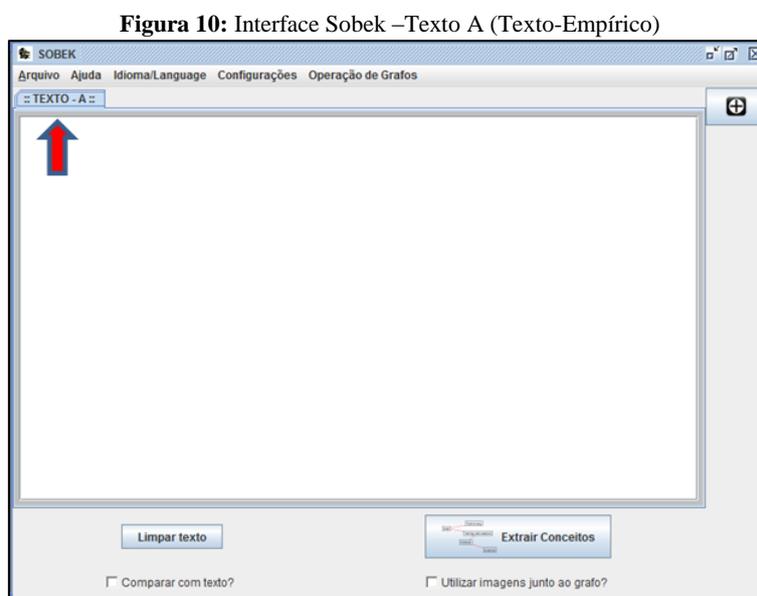
Portanto, a noção de panorama e padrão indica alguns estados da ação de linguagem que compõem o texto-empírico e, para tanto, interpretar com mais profundidade outros níveis de complexidade demandaria outras estratégias de apropriação da ação de linguagem do autor do texto-empírico, externa à ação de linguagem interpretativa do sujeito que minera um texto, para que esse possa arguir o que a utilização deste ou daquele signo visa. Frente a isso, a noção de recursividade da língua/linguagem faz-se presente, pois a atividade de produção textual exige um movimento recursivo de apropriação da linguagem de outro(s) para que a comunicação e os objetivos presentes na ação de linguagem (texto-empírico), em caráter arbitrário, possam fazer emergir no outro a condição de interpretante de mundos distintos ao seu, que tem por base motivações e intenções pessoais que visam a formas de validação da atividade de linguagem em práticas sociais. No próximo subcapítulo, apresenta-se o desenvolvimento de funções.

## **3.2 O desenvolvimento de funções**

Levando em consideração o estudo preliminar e, com o objetivo de desenvolver novas funções para a interface da ferramenta Sobek, contou-se com o apoio técnico do Grupo de Pesquisa em Tecnologia Aplicada à Educação (GTech-Edu), sob coordenação do Prof. Dr. Eliseo Berni Reategui, o qual é responsável pela manutenção da ferramenta Sobek. No entendimento de que esta tese não está pautada no estudo de códigos de programação, optou-se apenas por descrever o objetivo das funções elaboradas, as quais têm como finalidade apontar para níveis de estabilidade e variabilidade composicionais

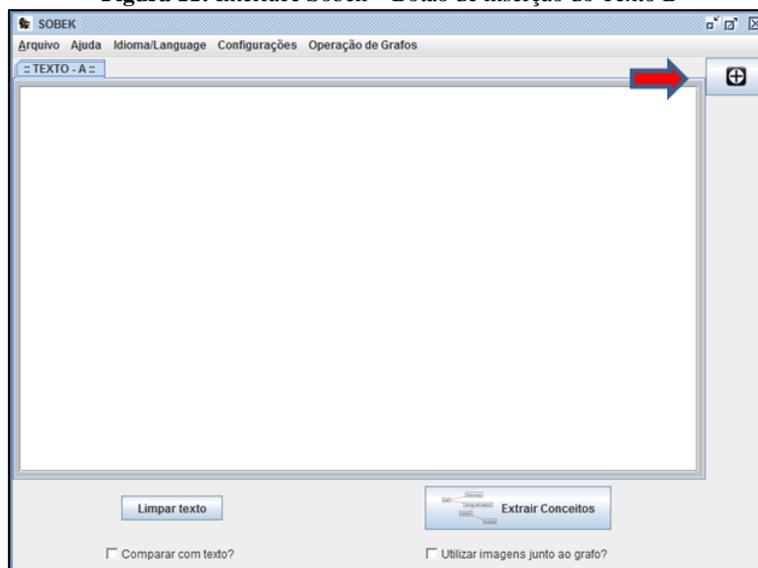
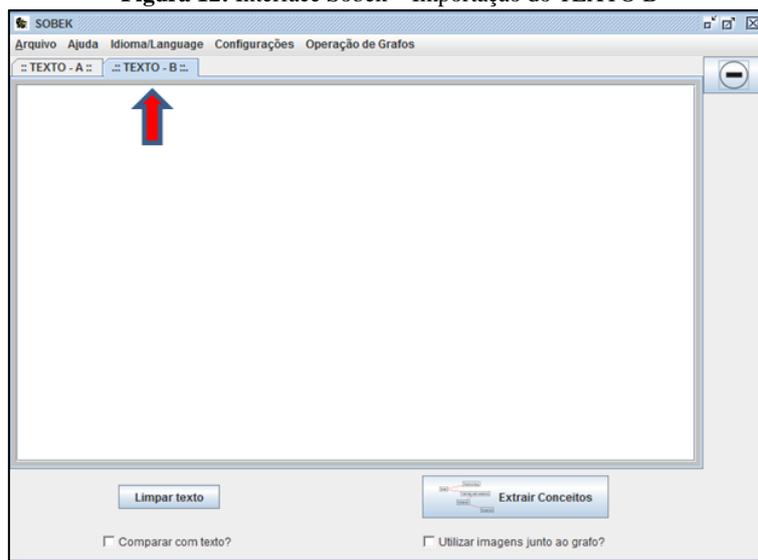
de textos gerados a partir de outros, já na forma de grafos distintos. A inserção de novas funções deu origem ao que fora chamado de nova versão da ferramenta Sobek, em Java, com vistas a ser, posteriormente, elaborada em HTML5, especificamente, dedicada à exploração das relações entre o conteúdo de dois textos, sendo esses oriundos de práticas e contextos empíricos de atividade de linguagem.

Conquanto no estudo preliminar anteriormente apresentado tenha-se percebido a necessidade de importar dois textos simultaneamente, para que o processo de mineração fosse totalmente automatizado, sem demandar uma ação humana para a identificação de conceitos estáveis ou variáveis em um texto-empírico, inseriu-se uma aba com a função intitulada TEXTO A, como pode ser visto na Figura 10, sendo esse o local de importação do texto-empírico, ou seja, produzido por um sujeito, a partir de outro texto-referência.



**Fonte:** Autor

A função TEXTO A permite operacionalizar o processo de mineração no formato da versão anterior, ou seja, realiza as mesmas ações da versão anterior, desde que não seja solicitada a mineração apenas desse texto. Caso o usuário tenha por finalidade comparar dois textos, então, por meio de um único clique no botão representado pelo sinal de + (mais) (Figura 11), ele poderá importá-lo automaticamente, e uma segunda aba, ao lado do botão Texto A, surgirá intitulada por TEXTO B (Figura 12), local destinado à importação do texto-referência.

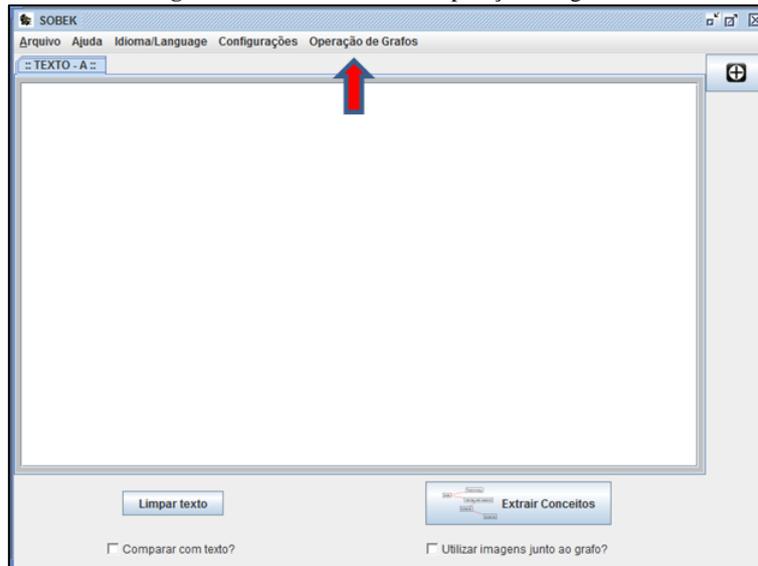
**Figura 11:** Interface Sobek – Botão de inserção do Texto B**Fonte:** Autor**Figura 12:** Interface Sobek – Importação do TEXTO B**Fonte:** Autor

Tendo em vista que foi inserida a possibilidade de inserção de dois textos simultaneamente, ação inexistente na versão anterior da ferramenta, entendeu-se que a extração de dados deveria permitir Operações de Grafos (Figura 13), com o objetivo de tornar visível mais de um panorama do todo da estrutura do texto.

A tomada de posição quanto à geração de tipos distintos de grafos é proveniente da concepção de que perspectivas panorâmicas diversas poderiam tornar elegíveis implicações relacionadas à complexidade do texto-empírico quanto à emergência de padrões plenamente associáveis à transição de um texto para outro. Para tanto, foram

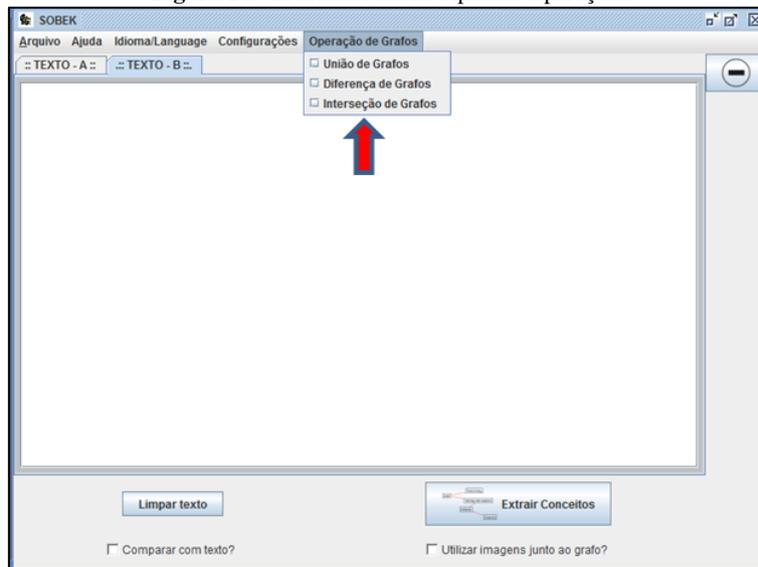
inseridas 3 opções de grafos (Figura 14), a saber: a) União de Grafos – a totalidade dos conceitos presentes, tanto Texto A (empírico) como Texto B (referência); b) Diferença entre Grafos – a extração de conceitos que estão no Texto A (empírico) e não estão no Texto B (referência), com o objetivo de evidenciar a variabilidade da ação de linguagem dos sujeitos, ao optar por determinados conceitos; c) Interseção de Grafos – conceitos que se repetem em ambos os textos, como visualização de padrões estáveis.

**Figura 13 - Interface Sobek: Operações de grafos**



**Fonte:** Autor

**Figura 14 - Interface Sobek: Tipos de Operações**

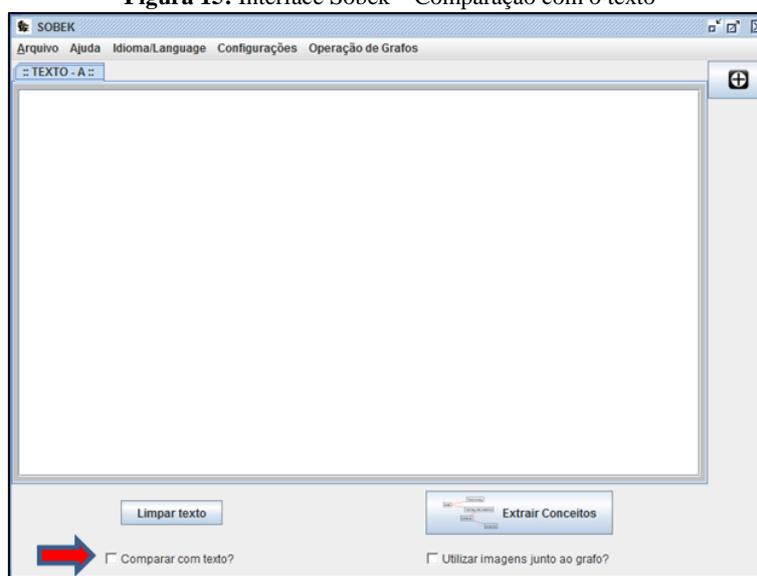


**Fonte:** Autor

Para, além disso, em um segundo momento, compreendeu-se que a opção de identificação de Diferença entre grafos exigia que a programação da ferramenta

operacionalizasse uma comparação não apenas entre o que era recorrente nos grafos dos textos referência e empírico, uma vez que comparar apenas os dados entre grafos, e não, da relação do conteúdo do grafo do texto-empírico à totalidade de dados do texto-referência propriamente dito permitiria que uma gama de dados pudesse passar despercebida pelo próprio processo de mineração. Para isso, programou-se a ferramenta a fim de que ela comparasse tanto as relações entre os grafos dos textos, como também todo o conteúdo do texto referência, tendo por base que a presença de uma palavra, apenas uma vez, deveria ser considerada, de forma a tornar o processo mais acurado. Referente a esse aspecto, foi inserida a função (botão) que pergunta se o usuário deseja *Comparar com o texto?* – como exposto na Figura 15. Essa estratégia está justificada pelo aporte da Teoria da complexidade que afirma que uma pequena alteração na constituição de um SAC pode representar mudanças significativas em sua constituição, assim como efeitos em relação a outros sistemas e subsistemas (LARSEN-FREEMAN, 1997; LARSEN-FREEMAN e CAMERON, 2008).

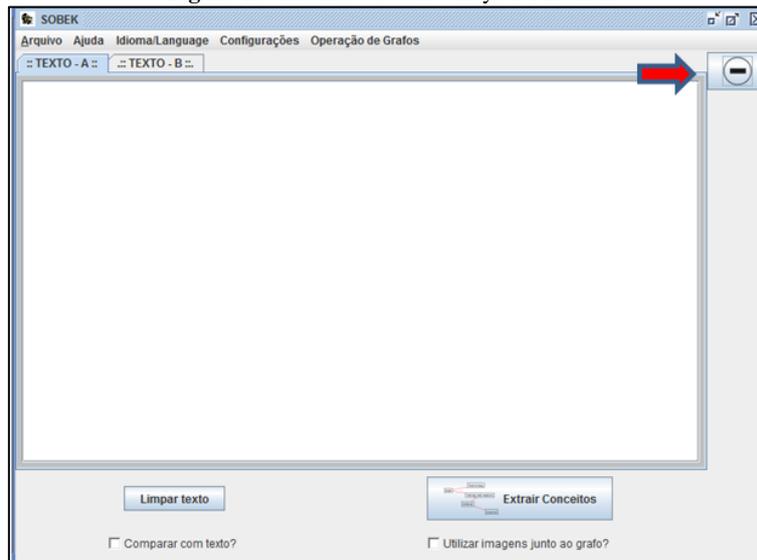
**Figura 15:** Interface Sobek – Comparação com o texto



**Fonte:** Autor

De modo a possibilitar o retorno navegacional da ferramenta à sua constituição em que se pode minerar apenas um texto, inseriu-se um botão, representado pelo sinal de – (menos) (Figura 16), com o qual se pode voltar ao *layout* anterior, sem considerar a importação de dois textos. Com esta alternativa, a ferramenta pode operacionalizar o processo de mineração de dados tanto no estilo anterior, quanto sob os moldes da nova versão.

**Figura 16:** Interface Sobek – Layout Anterior



**Fonte:** Autor

Uma vez apresentado o resultado do estudo preliminar que demandou a necessidade de desenvolvimento de uma nova versão da ferramenta Sobek, a seguir, apresenta-se o estudo final desta tese.

### 3.3 Estudo Final

Antes de apresentar a análise dos dados que compõem esta etapa do estudo, cabe discorrer acerca do posicionamento interdisciplinar (MORIN, 2008; LARSEN-FREEMAN e CAMERON, 2008; AXT, *et. al*, 2011) que emerge da base teórica da pesquisa quanto à atividade de linguagem, neste caso, a atividade de produção textual, a qual permitiu aos sujeitos colocarem em prática sua ação de linguagem. Nesse sentido, a presente análise de dados comporta duas dimensões: a primeira, de ordem teórica, ou seja, o encadeamento das proposições elencadas ao longo dos pressupostos teóricos; a segunda, de ordem aplicada, no que se refere à técnica de mineração de dados e seus desdobramentos, para fins de ampliação da perspectiva da Informática em relação ao conteúdo extraído de textos-empíricos. Relativamente à dimensão teórica, emerge deste estudo, em diálogo com as proposições do ISD e da Teoria da Complexidade, a ideia de que, quando da realização de uma atividade de produção que tem como condição inicial a interpretação de um texto-referência, tem-se, então, um texto externo à ação de

linguagem do sujeito que o interpreta (BRONCKART, 2006, 2007, 2008). Diante disso, é possível inferir que há uma tensão entre o conteúdo (motivações e intenções de seu(s) autor(es)), que é proveniente do texto-referência (externo) e a ação de linguagem do sujeito, aspecto que permite considerar, pelo menos, a existência de dois SACs em interação.

O texto-referência (externo) comporta em si outros sistemas e subsistemas que lhe conferiram a forma que apresenta, sendo ele oriundo de um escopo maior, que é o da atividade de linguagem (rede de conhecimentos que configuram os contextos interacionais) que, por sua vez, é a inscrição concreta das práticas coletivas e socialmente validadas à pretensão de verdade no mundo (BRONCKART, 2006, 2008). Nesse sentido, é possível retomar a concepção de que a atividade de linguagem é produtora de objetos de sentido ao atribuir valor semântico e semiótico (BRONCKART, 2008). Dessa forma, o texto-referência é uma manifestação de representações coletivas, de autoria de um sujeito específico, compostas por mecanismos de criação de unidades socioconvencionais (BRONCKART, 2008, p. 24): a) arbitrárias – por não se basearem em propriedades naturais e objetivas das entidades que exprimem; b) sociais – por estabilizarem as relações entre entidades verbais e seus referentes, pois determinam um acordo ou convenção entre membros de determinado grupo. Portanto, o texto-referência representa uma conexão a outros sistemas sociais, políticos e fins comunicacionais, não podendo ser mapeado tão facilmente, posto que ele é uma derivação de experiências dialógicas e polifônicas em caráter de recursividade.

Uma vez externo à ação de linguagem de um sujeito, o texto-referência, na perspectiva sistêmica da complexidade, instaura uma situação caótica de interação entre si e a ação de linguagem de seu leitor, a qual precisa ser materializada na forma de um texto-empírico escrito. Posto que o interesse desta tese está na transição (estado de fase) de um texto para outro (LARSEN-FREEMAN e CAMERON, 2008), o texto-referência está tomado como uma condição inicial para que o sujeito (aluno) possa tê-lo como força entrópica, que tensiona seus conhecimentos previamente internalizados, como forma de movê-los a outra zona do conhecimento (ZDR- ZDP – VYGOSTKY, 1998), propiciando que novos construtos sejam adquiridos na tensão gerada pela atividade de produção textual. Nessa perspectiva, têm-se dois sistemas distintos em questão. Portanto a atividade prescritiva de elaboração de um texto-empírico faz com que o interpretante tenha que acessar conhecimentos prévios à noção de textos, movimento esse que perturba as estruturas psíquicas já consolidadas, exigindo que ele assuma sua condição

de agentividade (BRONCKART, 2006; 2008). Desse modo, o texto-referência é um atrator, o qual exerce a instauração de uma relação entrópica à ação de linguagem individual, tendo em vista que a ação de linguagem do sujeito, quando da produção de um texto-empírico pode aderir ou não às propriedades (signos) presentes na estrutura do texto-referência. A influência exercida pelo texto-externo condiciona ação de linguagem a assumir plenamente sua condição de SAC, tornando-a (MORIN, 2008; LARSEN-FREEMAN e CAMERON, 1997, 2008; GLEICK, 1989) *complexa* e dinâmica, caótica e *adaptativa*, sensível às condições iniciais e aberta, recursiva, autossimilar e emergente, conforme exposição a seguir.

a) *Complexa* e dinâmica – por ter de dialogar com motivações e implicações outras, que constituem o texto-referência, de modo a extrair subsídios para a elaboração de um texto empírico; comporta um ecossistema de propriedades (signos) e sequências de dois sistemas, ou seja, do texto-referência e da ação de linguagem ainda não materializada na forma de um texto;

b) Caótica e *adaptativa* – devido a, possivelmente, ser composta por índices de estabilidade e variabilidade pouco previsíveis do conteúdo que lhe serviu de referência e, considerando-se que, aqui, a variabilidade é o ponto central desta pesquisa, por ser um aspecto tecnologicamente identificável pelo recorte produzido pela ferramenta Sobek; por interpretar os signos do texto-referência e, em um movimento discursivo não linear, adaptá-los às motivações e intenções do autor do texto-empírico, a fim de respeitar padrões socioculturalmente validados em relação à atividade de produção textual, ao contexto em que o texto-empírico circulará e assim por diante;

c) Sensível às condições iniciais e aberta – posto que o conteúdo empregado para a produção do texto-empírico pode ser estruturado por organizações similares ou completamente diferentes às propriedades do texto-referência, portanto, nunca dotada de características homônimas (LARSEN-FREEMAN e CAMERON, 1997, 2008; GLEICK, 1989) e porque precisa atribuir às propriedades (signos) do texto-referência significado e significante, ou seja, semiotizar seu conteúdo;

d) Recursiva – diante de um atrator estranho, frente ao caráter dialógico e polifônico que validam a ação de linguagem nos níveis individual e social, quanto ao tipo de texto pretendido pela atividade proposta ao sujeito e quanto ao conteúdo a ser interpretado, uma vez que as intenções do autor do texto-referência geram uma tensão entre mundos (mundo do autor e mundo do leitor);

e) Autossimilar e emergente – ao apresentar subdivisões do conteúdo do texto-referência, remetendo à noção de variabilidade mutacional (MANDELBROT, 1982) no que se refere ao caráter fractalizado acerca de determinados núcleos, informações reconfiguradas, na perspectiva do autor do texto-empírico, posto que, ao apresentar uma mudança de estado/fase, a qual tende a apresentar auto-organização quanto às estruturas que o compõem, neste caso, um texto-empírico, como resultante da transição de uma zona de conhecimento humano para outra. Essas zonas se enquadram na perspectiva discutida no subcapítulo dedicado à aproximação dos princípios de Vygostky à Teoria da Complexidade.

Ainda referente a este aspecto, entende-se, pela premissa interacionista sociodiscursiva, articulada pelo ISD, bem como pela caracterização dos SACs, abordada pela Teoria da Complexidade, que os estágios de desenvolvimento e de aprendizagem ocorrem simultaneamente, pois há uma interdependência entre o que é da dimensão do vivido (informalmente) e do aprendido (formalmente) (VYGOSTKY, 1998). Essa perspectiva se dá, também, pela compreensão de que as noções de coletivo e individual estão, indiscutivelmente, imbricadas na constituição do humano e das formas de consolidação do pensamento humano pela linguagem. Diante da concatenação das teorias apresentadas, cabe voltar à noção de interação entre o texto-referência e a ação de linguagem e o que pode ser observado por meio da ferramenta Sobek, no que diz respeito a uma ótica fractalizada. Ou seja, o entendimento de que a ferramenta de mineração possibilita a identificação de um fractal, dentre os tantos presentes no objeto de estudo desta pesquisa, faz com que a delimitação ocorra no nível da observação do que é estável ou variável em relação à forma do texto-empírico (ação de linguagem), no sentido mais literal da palavra, pois tratar do conteúdo demandaria outras abordagens teóricas e práticas que não estão aqui elencadas.

A geração de grafos via técnica de mineração, está arrolada como forma de identificar um padrão de ação de linguagem, relativo ao que parece estável, na transição de um sistema para outro, assim como o que supostamente emergiu dessa interação em condição de variabilidade do conteúdo do texto-referência. Cabe reforçar que a escolha pelas palavras *parece* e *supostamente* na redação do trecho anterior, deve-se à lúcida compreensão de que apenas pelo recorte da ferramenta de mineração não se pode mensurar, precisamente, a procedência da emergência dos signos capazes de serem evidenciados por essa técnica. Isso ocorre por duas razões: a primeira, porque todo signo linguístico é abstrato, imotivado e desdobrável; a segunda, porque o caráter

dialógico e recursivo da língua/ linguagem não permite “mapear” as especificidades das motivações e intenções do sujeito/autor apenas no nível de um signo, como feito por este estudo. Para uma concepção mais ampla quanto a isso, seria necessário observar o processo enunciativo dos sujeitos de pesquisa e propor entrevistas de verbalização das escolhas feitas por eles. Portanto, a aderência aos signos presentes no texto-referência, assim como sua forma e conteúdo, reconfigurados no texto-empírico, “nunca serão cópias idênticas das macroestruturas semânticas” (BRONCKART, 1999, p. 98) em relação ao texto que lhe serviu de condição inicial para a produção textual. A noção de condição inicial para determinada atividade pode ser vista como uma metáfora, frente à complexidade das experiências desenvolvimentais, anteriormente vividas pelos sujeitos, as quais estão sempre implicadas às ações que eles executam, pois constituem sua identidade e, como uma condição inicial não representa um “marco zero”, é apenas um ponto de partida.

A partir do que foi dito, concebe-se que o signo linguístico é um depositário das representações coletivas que só se consolida *na* e *pela* ação de linguagem, a qual situa seu valor representativo ou declarativo nas interpretações de mundo, razão por que nunca se torna totalmente estável frente à dinamicidade dos usos que os sujeitos lhe atribuem (BRONCKART, 2008). Nesse tocante, todo e qualquer signo é uma entidade arbitrária e desdobrável, fator esse que restringe a identificação precisa da utilização de um determinado signo e seu valor, como já mencionado. Ainda quanto à perspectiva arbitrária do que se pode minerar em um texto, compartilha-se da tese de Bakhtin (2006) que afirma que um signo possui natureza orgânica e móvel frente aos usos dados pelos sujeitos.

Logo, de antemão, o aporte teórico deste estudo permite dizer que os dados (signos) de um texto são, na verdade, em um primeiro momento, envelopes vazios (BRONCKART, 2006; 2008) que precisam ser “preenchidos” pela ação de linguagem dos sujeitos, com vistas a compreender o sentido que lhes fora dado. Desse modo, é preciso que o sujeito acesse mundos discursivos pertencentes a sua constituição identitária, sendo esses distintos do mundo ordinário de quem primeiramente o escreveu (autor do texto-referência), para que, ao atribuir sentido à ação de linguagem de outrem, possa aderir ao conteúdo de um texto externo ou refutá-lo à ação de produzir um texto-empírico.

A manifestação de dados, em caráter de estabilidade ou variabilidade de determinados signos, no texto-empírico, é o resultado da apropriação da ação de

linguagem do autor do texto-referência, o qual serve como instrumento/atrator que pode desencadear, ou não, a transição do sujeito, da Zona de Desenvolvimento Real (ZDR) para a Zona de Desenvolvimento Proximal (ZDP) (VYGOSTKY, 1998). Em realidade, a identificação de conservação (estabilidade) ou alteração (variabilidade) de signos em um texto-empírico pode apenas indicar tendências de um padrão de emergência da ação de linguagem, de modo que, frente à complexidade de um texto, como SAC, seriam necessários outros instrumentos de investigação para que se possa, concretamente, justificar as razões de um comportamento/padrão de emergência. Contudo, esta pesquisa se dispõe, a seguir, de maneira aplicada, a investigar de que forma a ferramenta Sobek pode evidenciar níveis de complexidade quanto aos padrões de estabilidade e variabilidade composicional do texto-empírico e sugerir a identificação de possíveis tendências de desenvolvimento (GLEICK, 1989; LARSEN-FREEMAN e CAMERON, 1997, 2008; MORIN, 2008).

Tendo em vista as postulações que emergiram do encontro das teorias adotadas nesta pesquisa, passou-se à dimensão prática e aplicada do estudo com a utilização da nova versão da ferramenta, considerando-se os aprimoramentos realizados referentes às novas funções que, aparentemente, condizem com a proposta de, além de gerar um panorama do todo da estrutura do texto-empírico, poder gerar grafos distintos.

No que tange à dimensão aplicada do estudo final, foram coletados textos que se classificam em dois grupos, a saber: a) textos-referência (artigo científico e capítulo de livro) e b) textos-produzidos (ou seja, textos-empíricos). Ambos são qualificados como gêneros textuais da esfera acadêmica (BRONCKART, 1999, 2008; MARCUSCHI, 2008) e de gênero discursivo secundário (BAKHTIN, 2003,2006), derivados de atividades de prática de produção textual. A pesquisa contou com as práticas da produção textual realizadas por duas turmas de nível de ensino superior, graduação em Artes Plásticas (rede pública) e Pedagogia (rede privada), ambas no estado do Rio Grande do Sul. A atividade envolveu a leitura extraclasse de um texto-referência, selecionado pelo(a) professor(a), para fins de produção de textos-empíricos, sendo esses todos produzidos por alunos (adultos entre 18-45anos), regularmente inscritos em turmas de nível de graduação. Portanto, os dados foram coletados das referidas turmas de graduação, situadas na área das ciências humanas, com o consentimento das instituições de ensino superior e dos respectivos professores de cada turma. Uma vez que os textos-empíricos, neste estudo, não são avaliados em sua totalidade discursiva, apenas a partir do ponto de vista do recorte da ferramenta Sobek, esses foram considerados válidos para

a finalidade da atividade proposta, a partir da ótica dos professores responsáveis pelas turmas.

Nesta etapa, realizou-se a coleta dos dados a serem submetidos ao processo analítico de mineração (CRESWELL, 2007) com a nova versão da ferramenta. Relativamente a isso, a pesquisa foi apresentada a dois professores responsáveis por duas turmas distintas, os quais, diante do interesse em ceder textos-empíricos produzidos por seus alunos em atividades de produção textual, comprometeram-se a enviá-los, via correio eletrônico, para o pesquisador desta tese. Os dados enviados deveriam conter o texto-referência, selecionado pelo professor como condição inicial para a atividade, e os textos-empíricos, produzidos pelos alunos sem qualquer observação avaliativa por parte dos professores. A participação dos professores e alunos se limitaria à produção dos textos e a seu envio sem qualquer outro tipo de envolvimento, uma vez que o objeto de estudo desta pesquisa é a ação de linguagem dos sujeitos (alunos), materializada na forma de texto-empírico.

Para tanto, o estudo final abarca dois textos-referência – um no formato capítulo de livro e outro no formato artigo científico – e 32 textos-empíricos produzidos pelo respectivo número de alunos, tendo cada discente redigido apenas um texto, totalizando, então, 32 produções. Metodologicamente, optou-se por intitular as turmas por Turma 1 e Turma 2, e os textos por T1, T2 e assim sucessivamente. Para fins de construção do *corpus* da tese, todos os sujeitos tomaram conhecimento dos objetivos da pesquisa, quando do recebimento do Termo de Consentimento Livre e Esclarecido (TCLE), e da vinculação dessa aos grupos de pesquisa N.E.S.T.A e GTech-Edu. No que tange às identidades e às produções textuais, essas foram mantidas em sigilo quanto à sua integralidade. Portanto, os dados coletados são apresentados na forma de fragmentos, de maneira a preservar a identidade dos sujeitos enquanto autores. O tratamento dos dados foi realizado pelo pesquisador quanto a submeter os textos à técnica de mineração de dados com a nova versão da ferramenta e analisá-los apoiado no aporte teórico do ISD e da Teoria da Complexidade. Essa delimitação se deu devido ao caráter inicial da proposição de novas funções, ou seja, considerou-se que a participação ativa de sujeitos, quanto ao uso da nova versão da ferramenta, seria uma abordagem investigativa posterior a este estudo, após o pesquisador ter constatado a sua real eficácia (da ferramenta) para a finalidade que se está objetivando a partir deste estudo.

Com referência à estratégia de exploração de níveis de complexidade (LARSEN-FREEMAN e CAMERON, 2008), a partir das definições teórico-

metodológicas desta tese, acerca dos dados/conceitos presentes nos textos, optou-se por apresentá-los, do seguinte modo:

- Turma (Turma 1, Turma2)
- Texto-referência com 3 grafos ilustrativos – a) nível 1 de complexidade (conceitos): propicia a visualização do padrão de recorrência no nível de conceitos, enquanto panorama mais geral da estrutura do texto; b) nível 2 de complexidade (inter-relações): contém um conceito selecionado para a visualização de inter-relações no nível de conceitos interconectados ou não, enquanto estrutura mais profunda do nível 1; c) nível 3 de complexidade (subinter-relações): inclui um conceito selecionado para a visualização do padrão de subinter-relações no nível de conceitos subinter-relacionados aos conceito de nível 2;
- Tabela com a descrição da totalidade de conceitos do texto-referência;
- Textos-empíricos – os dados minerados são apresentados na forma de 2 gráficos iniciais, o primeiro contendo o total de conceitos, inter-relações e subinter-relações, e o segundo, o total de conceitos e níveis de Intersecção (estabilidade) e Diferença (variabilidade), a partir de funções inseridas na ferramenta, sendo ilustrados a partir de 3 exemplos de grafos gerais, Intersecção e Diferença;
- Gráficos comparativos – 1) Nível 1 de complexidade – comparação entre o total de conceitos do texto-referência e do texto-empírico; 2) Nível 2 de complexidade – comparação entre o total de inter-relações do texto-referência e do texto-empírico; 3) Nível 3 de complexidade – comparação entre o total de subinter-relações do texto-referência e do texto-empírico.

A terminologia que denomina os níveis mencionados é um construto deste estudo, baseado na definição de SACs e nas postulações de Larsen-Freeman e Cameron (2008) e Capra (2002) e Morin (2008). Optou-se por descrever os conceitos no formato tabela com a finalidade de tornar visível a identificação feita de diferentes constituições, enquanto sistemas e subsistemas relacionados a cada conceito, inter-relação ou subinter-relação. A mesma abordagem foi conduzida quanto aos textos-empíricos, porém, para

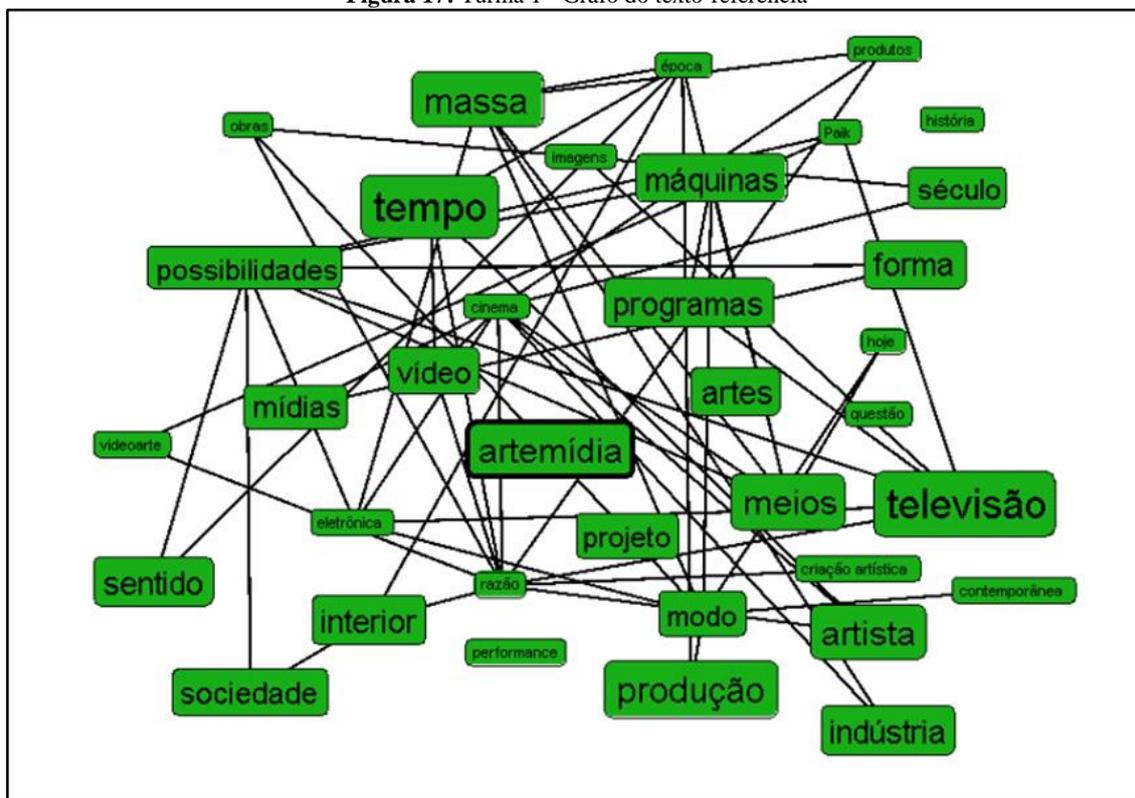
esses, apenas apresentam-se índices numéricos como forma de tornar factível a exibição da análise de todos os textos, do ponto de vista metodológico, conquanto o interesse central de análise está em evidenciar que se podem identificar níveis de complexidade distintos a partir da técnica de mineração de dados. Em relação à aplicação da nova versão da ferramenta aos textos-empíricos das duas turmas, esses, além de terem sido minerados sob a mesma lógica dos textos-referência, foram examinados quanto a suas funções de Intersecção entre grafos, para se poder visualizar o nível de estabilidade do conteúdo e o de Diferença para a variabilidade do conteúdo em relação ao texto-referência.

Uma vez que a função de grafo de União, mencionada na descrição das novas funções, permite a visualização de um único grafo composto pelo conteúdo dos textos referência e empírico, caracterizando o estado de ambos, esse não foi adotado, posto que o objetivo central, nessa etapa, era o de observar, isoladamente, conceitos estáveis ou variáveis com vistas a visualizar dados emergentes. Contudo, entende-se que essa função não é menos útil que as demais, pois ela admite a visualização de todas as propriedades dos sistemas envolvidos (textos) na atividade de produção textual (BRONCKART, 2006, 2007, 2008), aspecto que permite ao minerador/professor uma leitura automatizada do todo composicional de ambos os textos (ação de linguagem do autor externo e do autor atuante). Os gráficos que demonstram a comparação entre os índices dos textos-referências e textos-empíricos são meramente ilustrativos, posto que, como antes defendido, cada texto configura-se como um sistema distinto, porém, neste caso, sendo o texto-referência a condição inicial para a elaboração de um texto-empírico, considerou-se relevante visualizar a não linearidade da ação dos sujeitos em relação ao que havia sido lido.

Na Turma 1 o texto-referência selecionado pela professora para a atividade de produção textual foi um capítulo de livro, com a temática a respeito de Artemídia, de relativa extensão (20 páginas), o que suscita um número elevado de conceitos (signos) em sua estrutura e encadeamento de ideias. Após a leitura do texto-referência, em caráter de atividade extraclasse, os alunos deviam produzir seus textos empíricos. O total de participantes nesta turma foi de 13 alunos, os quais geraram, respectivamente, 13 textos empíricos. Diante do recebimento dos textos da Turma 1, primeiramente, tratou-se de submeter o texto-referência (Figura17) ao proceso de mineração de dados com a ferramenta Sobek, em sua nova versão, porém, ainda no estilo da versão anterior, ou seja, com a finalidade de gerar um grafo composto apenas pelo conteúdo desse texto,

sem relacioná-lo a qualquer texto-empírico, com vistas a explorar detalhadamente sua constituição. No que tange ao tratamento dos dados, do conteúdo apresentado nos grafos, durante toda a realização dos processos de mineração, selecionou-se a opção de extração de 50 conceitos por grafo, bem como foram refutados artigos definidos, indefinidos e preposições. Dessa maneira, na Figura 17, apresenta-se um panorama geral dos conceitos mais recorrentes no texto-referência, entendendo-o como enquadrado no Nível 1 de complexidade, uma vez que revela uma macroestrutura do texto composta pela totalidade de seus dados/conceitos e relações. Essa macroestrutura em si oferece uma perspectiva mais geral de um estado composicional do texto, evidenciando um primeiro padrão/estado interno desse, pautado pelo caráter exclusivo de repetição de signos linguísticos adotados *na* e *pela* ação de linguagem do autor do texto, externo à ação de linguagem de seu leitor (BRONCKART, 2006, 2007, 2008).

**Figura 17:** Turma 1 - Grafo do texto-referência

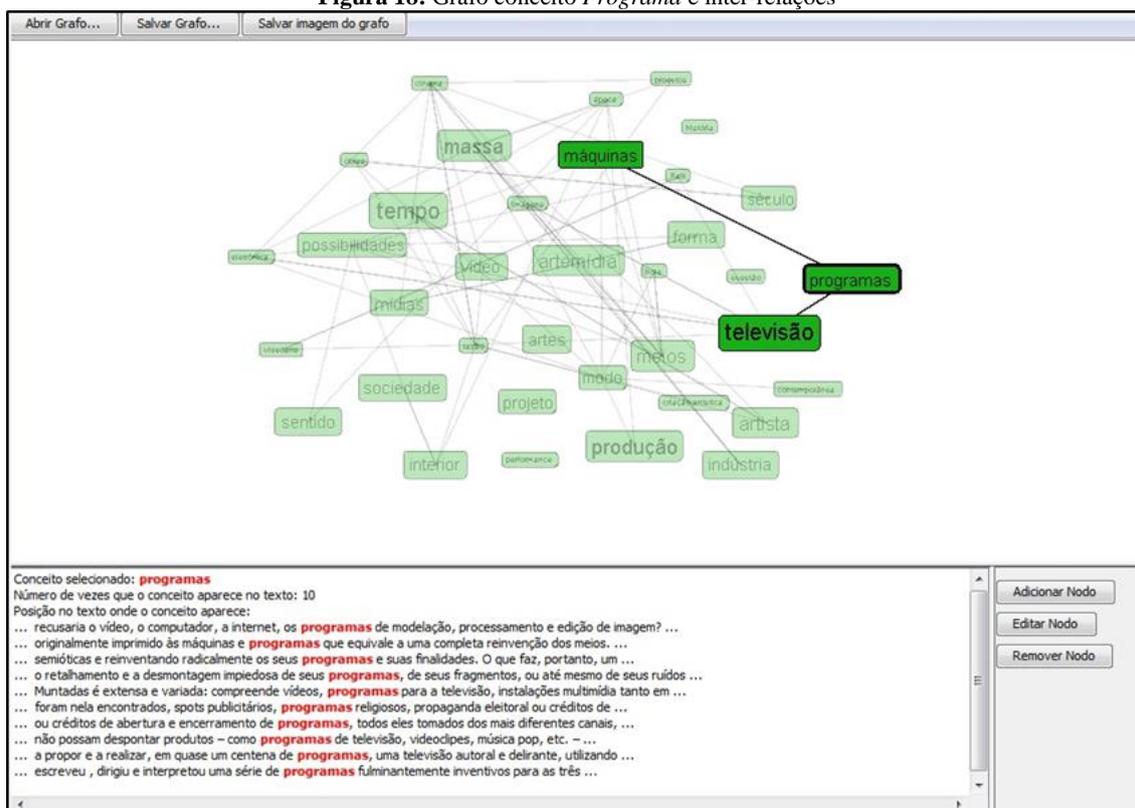


Fonte: Autor

De maneira a aprofundar a perspectiva complexa do processo de extração e interpretação dos dados do texto referência, em seguida, observou-se a existência do Nível 2 de complexidade, pelo qual se pode aprofundar a dimensão do Nível 1, na medida em que este permite analisar, de modo fractalizado (LARSEN-FREEMAN e CAMERON, 2008, GLEICK, 1989), outra dimensão da ação de linguagem do autor,

entendo-a como um SAC, que em si comporta múltiplos sistemas e subsistemas. Esta operacionalização é possível ao clicar em um conceito (nodo), o qual passa a exibir as inter-relações associadas ao primeiro nível exibido na Figura 17. Para ilustrar essa ação, selecionou-se o conceito *Programas*, o que permitiu observar sua inter-relação aos conceitos máquinas e televisão (Figura 18), assim como o usuário/tutor/profesor pode analisar abaixo a posição deste no texto e o número de vezes que foi recorrente, embora essa ação complementar não esteja em questão no estudo final.

**Figura 18:** Grafo conceito *Programa* e inter-relações



Fonte: Autor

Já no que se refere ao Nível 3 de complexidade, ele está compreendido como nível de visualização de subinter-relações, em que se pode obter uma perspectiva ainda detalhada da ação de linguagem de quem produz um texto, por meio de associação relacionada a inter-relações e relações. Portanto, é possível realizar, visualmente, uma prospecção e retrospectão dos caminhos da ação de linguagem do autor do texto, seja ele referência ou empírico. Para ilustrar o nível 3, o conceito *Possibilidades* foi selecionado e, a partir dele, pode-se analisar uma camada ainda mais fractalizada, que diz respeito à presença de suas subinter-relações. Na Figura 19, pode-se notar duas inter-relações (sentido e sociedade) e seis subinter-relações (eletrônica, Paik, meios, máquinas, televisão e forma) vinculadas ao conceito *Possibilidades*. Essa estratégia de



texto referência, em um primeiro momento, foi necessário minerar e explorar o conteúdo gerado pela ferramenta, na forma de grafos, para que fosse possível compará-los aos dados dos textos-empíricos acerca de seus níveis de estabilidade e variabilidade composicional (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008).

**Tabela 2:** Níveis de complexidade do Texto-referência (Turma 1)

| <b>Nível 1<br/>(Conceito)</b> | <b>Nível 2<br/>(inter-relação)</b>                       | <b>Nível 3<br/>(subinter-relação)</b>        |
|-------------------------------|--|--|
| 1. Artemídia                  | inexistente  | inexistente                                  |
| 2. Artes                      | inexistente  | inexistente                                  |
| 3. Artista                    | cinema   | tempo, razão                                 |
| 4. Cinema                     | eletrônica, vídeo, indústria, meios, mídias              | razão, produtos, século, artista             |
| 5. Contemporânea              | modo   | inexistente                                  |
| 6. Criação Artística          | razão  | inexistente                                  |
| 7. Eletrônica                 | massa  | possibilidades, televisão, massa, modo       |
| 8. Época                      | tempo, sentido, interior, produção                       | massa, meio                                  |
| 9. Forma                      | possibilidades, mídias                                   | inexistente                                  |
| 10. História                  | inexistente  | inexistente                                  |
| 11. Hoje                      | modo, meios  | inexistente                                  |
| 12. Imagens                   | televisão  | inexistente                                  |
| 13. Indústria                 | massa, cinema  | inexistente                                  |
| 14. Interior                  | sociedade, razão, época                                  | inexistente                                  |
| 15. Máquinas                  | programas, modo  | possibilidades, meios                        |
| 16. Massa                     | produto, indústria                                       | meios, época, eletrônica, modo               |
| 17. Meios                     | hoje, cinema   | possibilidades, máquinas, época, massa       |
| 18. Mídias                    | cinema, forma  | inexistente                                  |
| 19. Modo                      | produção, contemporâneo, hoje, máquinas, obras           | eletrônica, massa                            |
| 20. Obras                     | razão, século, modo                                      | inexistente                                  |
| 21. Paik                      | videoarte  | televisão, possibilidades                    |
| 22. Performance               | inexistente  | inexistente                                  |
| 23. Possibilidades            | sentido, sociedade, forma                                | eletrônica, Paik, televisão, meios, máquinas |
| 24. Produção                  | modo, época  | inexistente                                  |
| 25. Produtos                  | massa  | cinema, razão                                |
| 26. Programas                 | televisão, máquinas                                      | inexistente                                  |
| 27. Projeto                   | inexistente  | inexistente                                  |
| 28. Questão                   | inexistente  | inexistente                                  |
| 29. Razão                     | vídeoarte, interior, criação artística, televisão, obras | cinema, produtos, artista, vídeo, tempo      |
| 30. Século                    | cinema, obras  | inexistente                                  |
| 31. Sentido                   | possibilidades, época                                    | inexistente                                  |
| 32. Sociedade                 | interior, possibilidades                                 | inexistente                                  |
| 33. Televisão                 | programas, razão   | Paik, eletrônica, possibilidades, imagens    |
| 34. Tempo                     | época  | razão, vídeo, artista                        |
| 35. Vídeo                     | cinema   | tempo, razão                                 |
| 36. Videoarte                 | Paik, razão  | inexistente                                  |

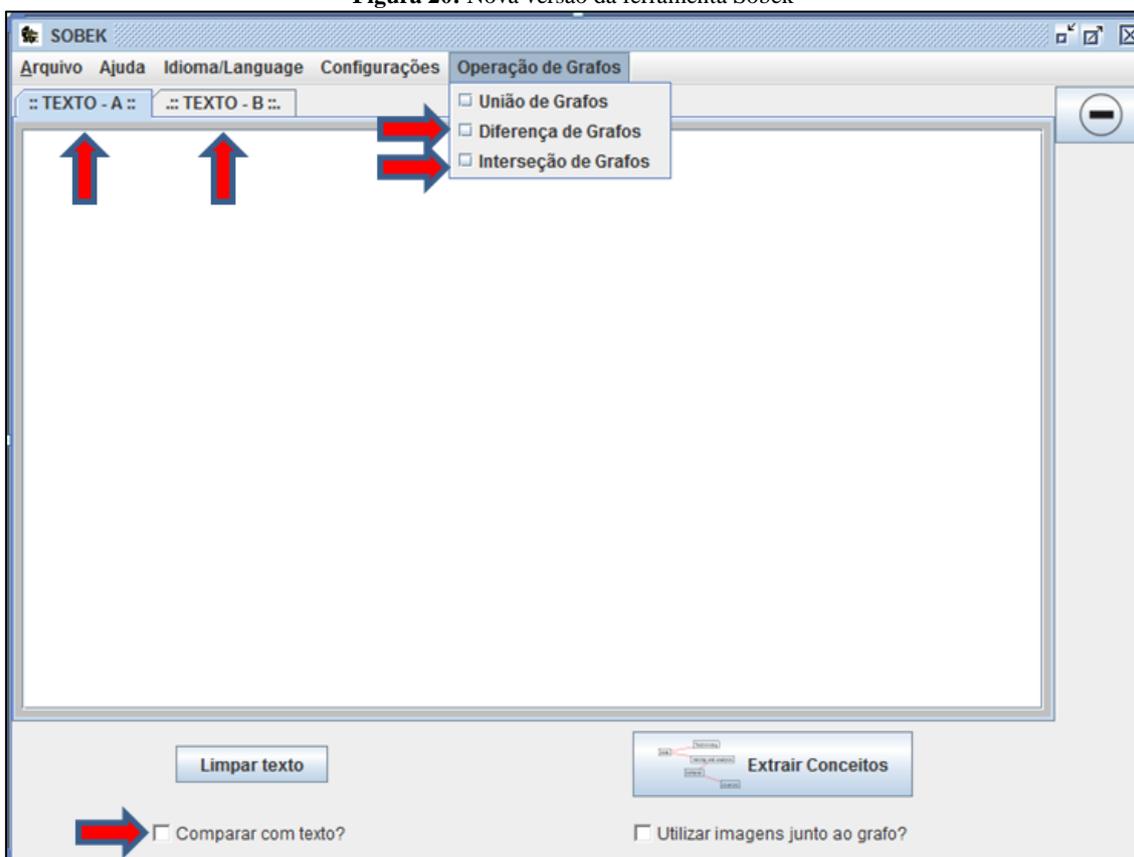
Fonte: Autor

A análise do texto-referência da Turma 1 permitiu identificar um total de 36 conceitos, 66 inter-relações e 47 subinter-relações. Cada conceito apresenta núcleos distintos (sistemas e subsistemas) de inter-relações e subinter-relações, como visto na Tabela 3, aspecto que aponta para a eficiência de Sobek ao tornar evidente diferentes perspectivas para cada nível de complexidade do texto. Contudo, ressalta-se que se está considerando que os conceitos, igualmente utilizados nos textos-empíricos, sejam apenas repetições gráficas dos conceitos presentes no texto-referência, pois, como antes mencionado, os textos-empíricos e os conceitos não são cópias idênticas dos conceitos e do texto-referência (BRONCKART, 2006; 2008). Essa interpretação também abarca a inferência de que os signos linguísticos, mesmo que tomados por “padrão” no nível coletivo, não têm a mesma carga de produção de sentido para diferentes sujeitos. Ou seja, a adesão a este ou a aquele signo está sempre pautada pela arbitrariedade semântica que o signo linguístico carrega em si, seja em sua aplicação isolada ou situada em uma estrutura sintática.

Em relação ao texto-referência da Turma 1, o usuário/professor pode, previamente à realização da atividade de produção textual, fazer uso da ferramenta e ter em mente os caminhos percorridos pelo autor do texto, para que, *a posteriori*, possa relacioná-los aos presentes nos textos-empíricos de seus alunos. Essa estratégia pode auxiliar o docente a perceber a necessidade de tratar de outros conceitos e/ou temas pertinentes à atividade que possam conduzir os alunos a uma produção textual consciente da relevância de alguns pontos a serem discutidos (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008). O discurso do usuário/professor pode gerar subsídios para que, entrópicamente, a influência do texto-referência possa gerar uma transição entre zonas do conhecimento, ou seja, para que o conteúdo permita a ampliação do pensamento consciente acerca de determinados sistemas e subsistemas presentes no texto, conduzindo o leitor e produtor do texto-empírico de sua Zona de Desenvolvimento Real à Zona de Desenvolvimento Proximal, prevista pelo professor. Essa transição entre zonas está subentendida conforme discutido anteriormente, no sentido de que a apropriação linguística de signos e sua dimensão significado-significante seja consolidada do social ao individual, permitindo que os sujeitos “migrem” de um estado de fase (zonas) para outro, pela mediação de outro sujeito, textos, etc.

Para a mineração dos textos-empíricos, utilizou-se a nova versão da ferramenta (Figura 20) para a geração de grafos distintos: a) um grafo geral do conteúdo recorrente, isoladamente, para que conceitos, inter-relações e subinter-relações fossem registrados; b) um grafo com a operação de Intersecção (estabilidade) para verificação do percentual de “repetição” de conceitos; c) um grafo com a operação de Diferença (variabilidade) para apontar possíveis emergências sistêmicas, sendo as ações b) e c) realizadas individualmente e, sempre, levando em consideração a comparação automatizada do conteúdo do texto-empírico, não apenas ao grafo do texto-referência, mas também à integralidade.

Figura 20: Nova versão da ferramenta Sobek



Fonte: Autor

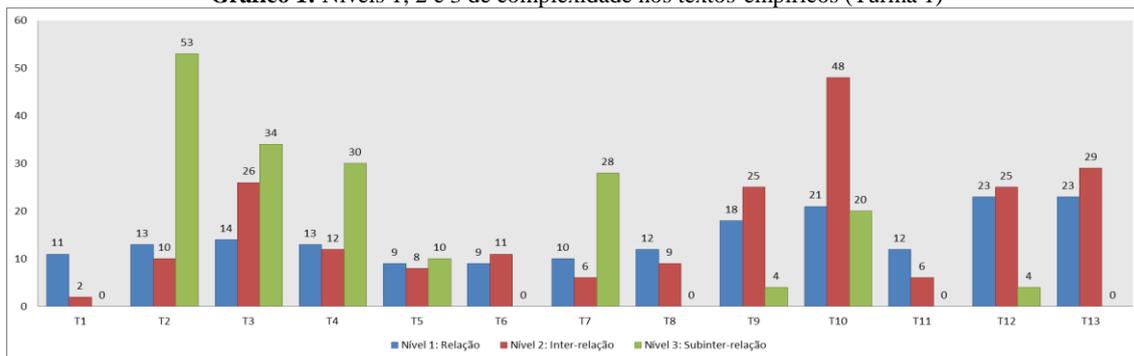
A partir da estratégia metodológica de mineração adotada para analisar, por meio de grafos, o total de 13 textos produzidos na Turma 1, foi possível identificar níveis de complexidade distintos entre os textos-empíricos (GLEICK, 1989; LARSEN-FREEMAN e CAMERON, 1997, 2008; MORIN, 2008), aspecto que corrobora que a ação de linguagem dos sujeitos se configura de modo complexo, dinâmico, adaptativo a suas motivações e intenções que os levaram à construção de seus textos. Dessa forma, posteriormente ao processo de extração de dados, observou-se que, por mais que, no

conteúdo escrito, forma e finalidade fossem semelhantes, os textos-empíricos, quando minerados acerca de seus conceitos, inter-relações e subinter-relações, mostram-se completamente singulares, tendo sido possível identificar que cada autor/aluno percorreu, por meio de sua ação de linguagem, trajetórias distintas.

No Gráfico 1 podem-se visualizar os percentuais observados em cada texto, o que revela que os textos produzidos na atividade de produção textual possuem índices distintos em cada nível e que alguns se sobressaem em relação aos demais, seja por índices inferiores e/ou superiores a eles. Se observados sob o viés do Nível 1 de complexidade (conceitos), tem-se, em geral, um número de conceitos bem próximo, oscilando entre 9 e 14 conceitos, assim como em outros textos percebeu-se uma variação entre 18 e 23 conceitos.

No que tange ao Nível 2 de complexidade (inter-relações), nota-se que há índices bem superiores aos do primeiro nível, que variaram entre 2 e 48, demonstrando que a ótica adotada permitiu evidenciar a complexidade dos textos em níveis distintos quanto à ação de linguagem dos sujeitos. Quanto ao Nível 3 de complexidade (subinter-relações), percebeu-se que nesse os índices oscilaram entre 0 (inexistente) e 53 conceitos, fator que reforça o posicionamento de que, embora a ferramenta gere um panorama do texto minerado, cabe, sem dúvida, investigar outros níveis em caráter de rede que, em um primeiro momento, se observado apenas o grafo geral (Nível 1), podem parecer inexistentes, mas, na verdade, estão situados em outras camadas do SAC.

Os índices apontados em cada nível indicaram que a ação de linguagem dos sujeitos não pode ser tomada como linear, por mais que pareça sê-lo no formato texto escrito, pois, com a utilização da ferramenta, sob o viés proposto, pode-se notar a real complexidade presente na atividade de linguagem em que os sujeitos estavam engajados (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008). Além disso, cabe ressaltar que, embora os textos-empíricos tenham tido o mesmo texto-referência como condição inicial para a produção textual, por meio da ferramenta de mineração de dados, observou-se que os níveis de complexidade são distintos entre si, seja no nível 1, 2, ou 3 de complexidade sistêmica.

**Gráfico 1:** Níveis 1, 2 e 3 de complexidade nos textos-empíricos (Turma 1)

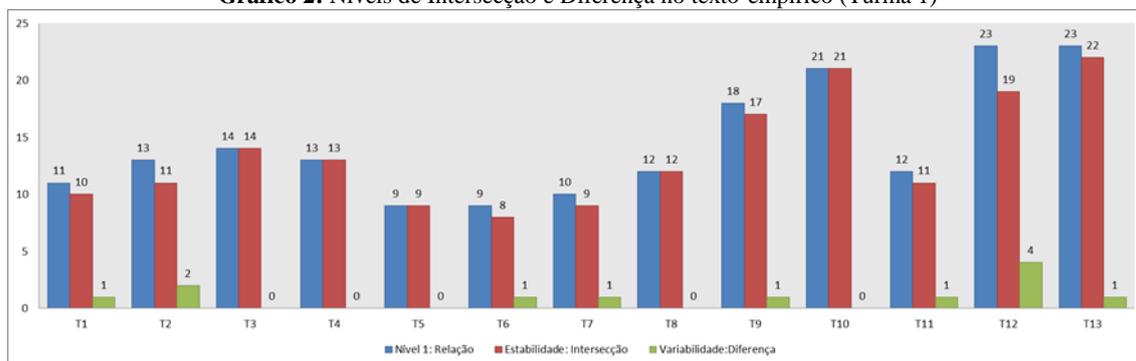
**Fonte:** Autor

Ao tornar tecnologicamente observável a ação de linguagem enquanto texto, a partir da utilização da ferramenta Sobek, infere-se que os textos-empíricos mostraram-se, em geral, mais complexos na proposição dos níveis 2 e 3. Embora alguns textos não apresentem subinter-relações, como é o caso dos textos T1, T6, T8, T11 e T13, não se pode justificar tal ausência apenas pela extensão do texto, tendo em vista, por exemplo, que o texto T13 possui número significativo de conceitos e inter-relações, mas não apresenta níveis de subinter-relações. Em outras palavras, a partir da proposição da leitura de um texto-referência em uma atividade de produção textual, não se pode prospectar concretamente a complexidade presente em um texto-empírico, posto que é a ação de linguagem (o sujeito/autor) quem se encarregará de complexificar ou descomplexificar o que emerge dessa prática.

Ainda no que se refere ao Gráfico 1, entende-se que os índices apontados, se observados, podem fazer com que um professor, por exemplo, exerça sua mediação pedagógica, para fins de atingir o objetivo que deseja na atividade, socilitando, por exemplo, a reescrita de textos, a partir da inserção de outros conceitos, tendo em mente essas variáveis coletivas enquanto panorama geral do desempenho da turma acerca de sua articulação acompanhada entre o texto-referência e o texto-produzido. Com a finalidade de ilustrar a análise realizada a partir da inserção das funções de Intersecção, denominada aqui como conceitos em estabilidade e Diferença, entendida como conceitos em varibilidade, sendo elas o cerne deste estudo, na medida em que se pode visualizar, a partir delas, as turbulências/transições/emergências da ação de linguagem dos sujeitos, no Gráfico 2 têm-se, primeiramente, o total de conceitos manifestados pelos sujeitos e, em seguida, o índice de Intersecção, ou seja, conceitos que se mantiveram em estabilidade ao conteúdo do texto-referência e de Diferença, sendo esses

aqueles conceitos que representam a variabilidade composicional do texto-empírico em relação ao texto lido.

**Gráfico 2:** Níveis de Intersecção e Diferença no texto-empírico (Turma 1)



Fonte: Autor

No intuito de investigar uma perspectiva ainda mais complexa, além dos 3 níveis de complexidade evidenciados (Gráfico 1), no Gráfico 2, pode-se notar que os 13 textos-empíricos da Turma 1 têm índices distintos de conceitos e, de modo geral, apresentaram um padrão de estabilidade ao conteúdo do texto-referência. Os índices deixam claro que, embora alguns textos tenham número maior de conceitos gerais, os níveis de conceitos em estabilidade, minerados na condição de intersecção (conceitos que estavam no texto-referência e permaneceram no texto-empírico), são, em todos os textos, radicalmente superiores aos níveis de variabilidade, ou seja, de conceitos diferentes que tenham emergido da prática de leitura e produção textual. Por sua vez, esse aspecto torna visível uma ação de linguagem atrelada ao conteúdo do texto-referência e que, por fatores desconhecidos, tendem a fazer com que o sujeito/autor mantenha as propriedades que o influenciaram à elaboração de seu texto (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008).

O padrão de estabilidade (LARSEN-FREEMAN e CAMERON, 2008) remete, portanto, a uma constatação que emerge desse estudo: a de que os alunos, na maioria das vezes, preferem manter a estrutura geral do texto-referência sem arriscar novas frentes de argumentação, narração e/ou descrição, talvez, por considerarem que assim não se distanciariam do tema central em questão no texto. Por outro lado, sob o viés da complexidade, considera-se que o nível de emergência da ação de linguagem dos sujeitos dessa turma possui baixo nível de conceitos emergentes, havendo mais auto-organização no sentido de se preparar para conservar, senão a forma, o conteúdo do texto-referência. Ressalta-se que, muito embora os alunos tenham tido certa liberdade autoral na produção de seus textos, essa limitou-se à conservação dos conceitos do texto

lido. Alguns textos, por exemplo, como os textos T3, T4, T5, T8 e T10, sequer contêm índices de conceitos diferentes dos lidos, aspecto que demonstra total conservação da ação de linguagem do autor do texto referência, mesmo considerando-se que o autor/aluno semiotiza os conceitos que apresenta em seu texto, de modo a adaptá-lo dentro da organização que dá a ele. É interessante observar que não apenas textos curtos, como o T5, não apresentam variabilidade conceitual, ao passo que o T10, de extensão significativa, também não possui tal variação. Nesse sentido, o que se entende é que o processo minerativo demanda, necessariamente, mais de uma estratégia para poder, de fato, manifestar a complexidade de um texto.

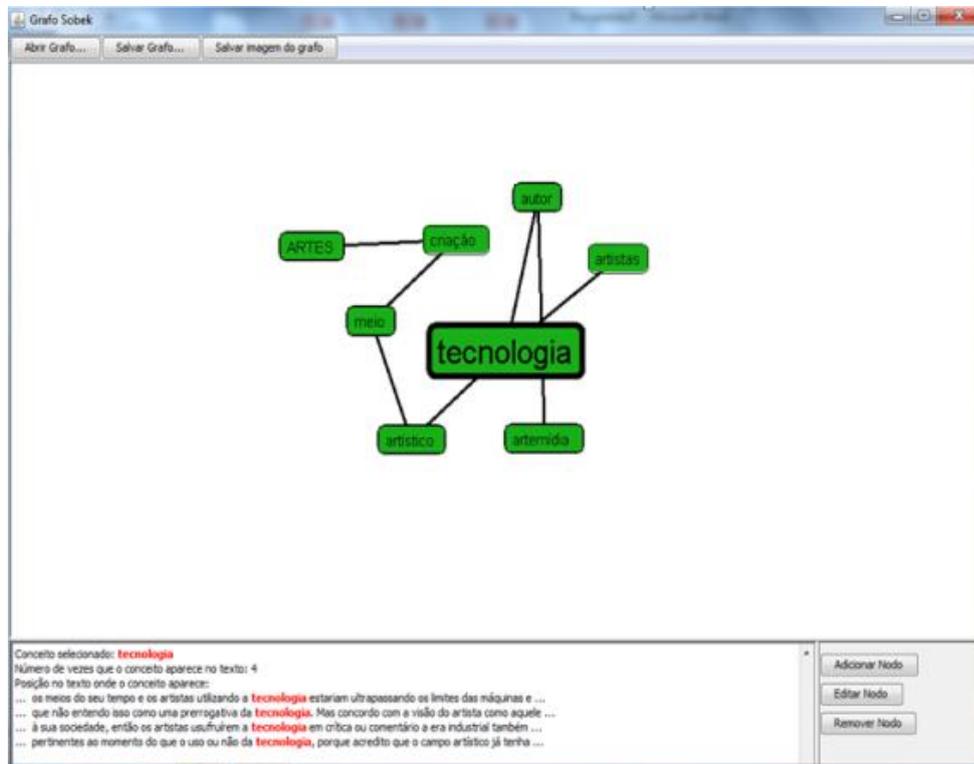
A identificação de uma variável coletiva em padrão de estabilidade permite que o usuário/professor possa se articular em relação à mediação que considera executável em relação à ação de linguagem de seus alunos, uma vez que, a partir desse padrão, ele poderá rever suas ações, refletir acerca das zonas de desenvolvimento de seus alunos, no sentido de tentar propiciar a inserção de estratégias que gerem a transição da Zona de desenvolvimento Real para a Zona de Desenvolvimento Proximal (VYGOSTKY, 1998), entendendo que, quanto mais variações os textos apresentarem, maior, supostamente, será o grau de apropriação da ação de linguagem dos sujeitos em relação ao texto lido, bem como maior será a emergência de novas condições de interação e aprendizagem.

Com a finalidade de ilustrar a ideia de que, embora haja um padrão de estabilidade, observado em uma dimensão mais profunda do processo de mineração e diante da inserção das novas funções de operação entre grafos, já que há uma diversidade composicional nos textos, optou-se por apresentar grafos específicos dessas questões: a) três grafos que mostram a dinamicidade da ação de linguagem dos sujeitos em perspectiva de índices distintos; b) três grafos que mostram a aplicação e identificação da função de Intersecção (estabilidade); c) três grafos com a função de Diferença (variabilidade).

As operações entre textos, como é o caso das funções de Intersecção e Diferença, não possuem ligações para a maioria dos conceitos, uma vez que são analisados pela ferramenta de modo isolado, porém acredita-se que esse aspecto em nada afete a proposição e a aplicabilidade dessas operações. Em se tratando da dinamicidade e auto-organização em padrão de estabilidade nos textos-empíricos, o índice de conceitos oscila e representa o grau de adaptabilidade dos conceitos do texto lido. Deste modo, no T6 (Figura 21), nota-se um índice baixo de conceitos em relação

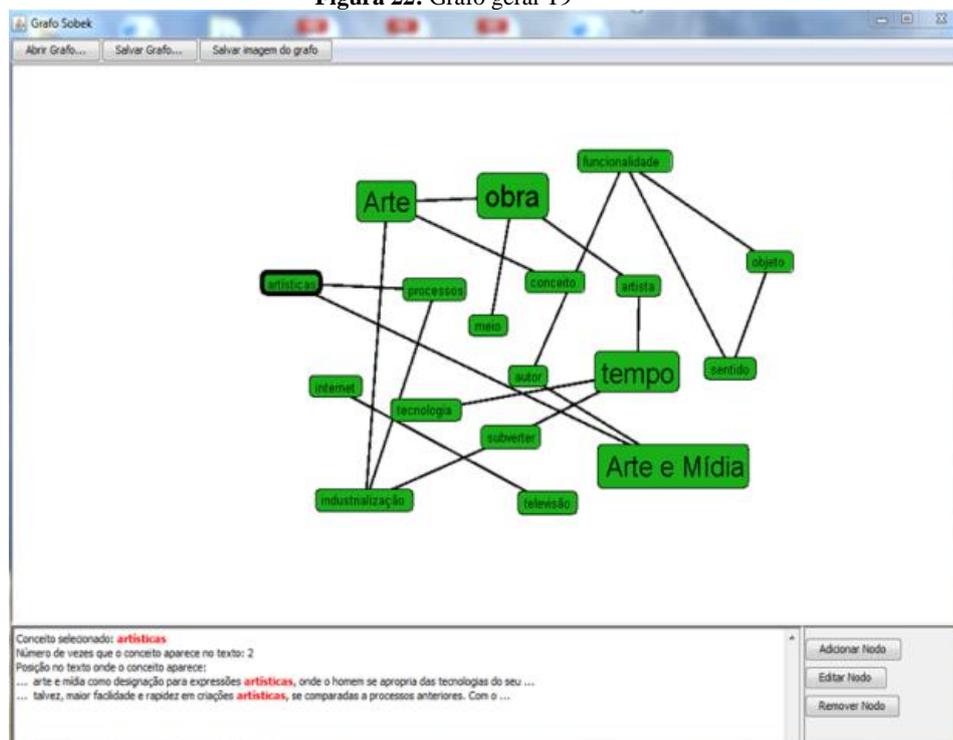
aos demais, ao passo que o T9 (Figura 22) contém um índice médio de conceitos, seguindo-se o T13 (Figura 23) que possui um índice alto de conceitos.

**Figura 21:** Grafo geral T6



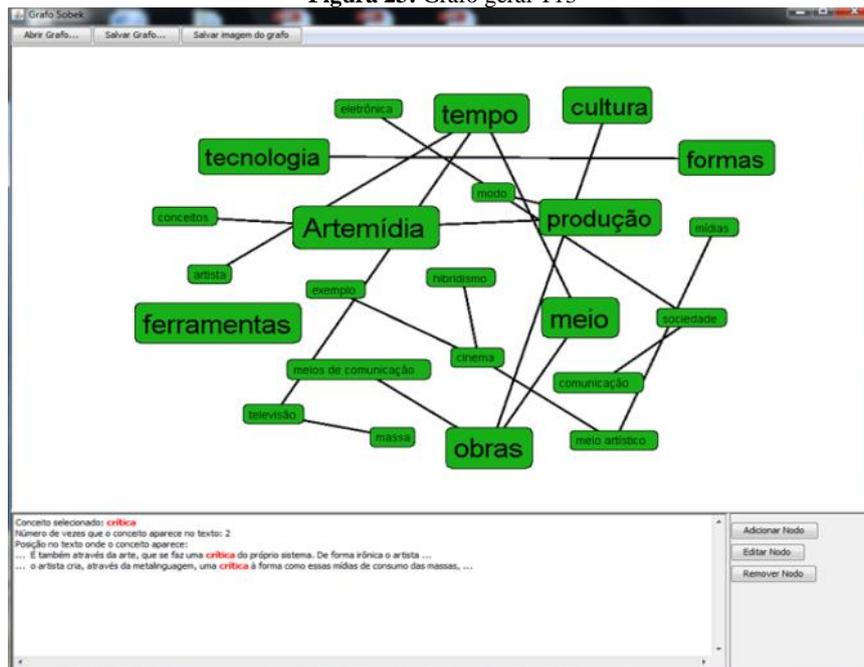
Fonte: Autor

**Figura 22:** Grafo geral T9



Fonte: Autor

Figura 23: Grafo geral T13



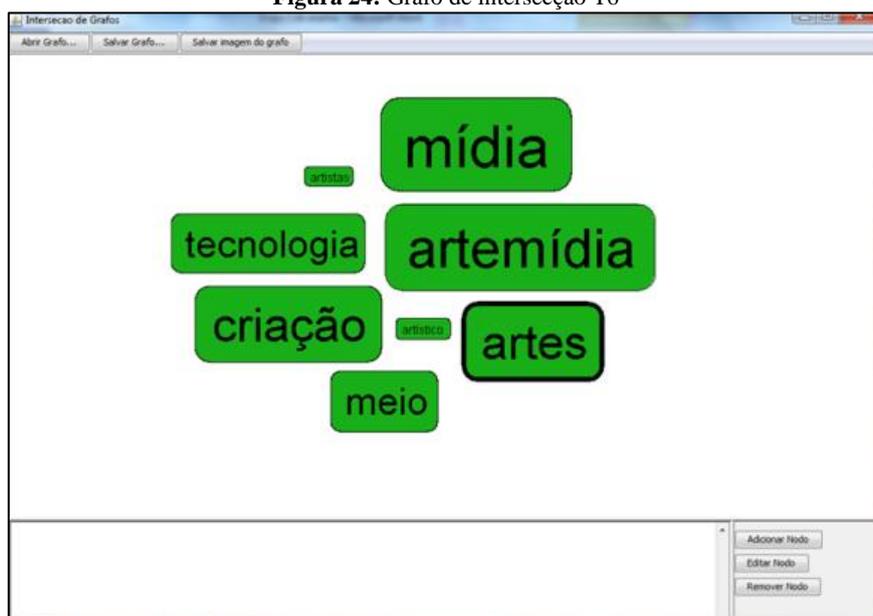
Fonte: Autor

A visualização da composição geral dos textos T6, T9 e T13 (Figuras 21, 22 e 23) permite afirmar, mesmo tendo, supostamente, a mesma orientação para a produção de um texto empírico, que a ação de linguagem dos sujeitos enquanto SAC tende a manifestar características distintas quanto à dinamicidade presente na elaboração de um texto-empírico, que demanda do sujeito/autor que ele acesse suas experiências prévias (desenvolvimento) e os processos de aprendizagem aos quais já foi submetido, o que reforça a afirmação de Vygostky (1998) de que aprendizagem e desenvolvimento são estados de fase, ou seja, SACs, que interagem simultaneamente, e não, em caráter isolado. Com isso, se está afirmando que os sujeitos transcendem a proposição da tarefa, enquanto condição inicial para elaborar um texto e utilizam-se de fontes que se consolidaram na base do pensamento humano, por meio de sua ação no mundo (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008). Portanto, os três grafos apresentados podem ser considerados subsistemas que se constituem não só pela ação de linguagem do autor do texto-referência, mas também, pelas motivações e intenções dos autores de cada texto-empírico. Já no que diz respeito aos índices de inter-relações e subinter-relações, esses podem ser vistos no Gráfico 1, bem como a indicação de níveis de estabilidade ou variabilidade é mostrada no Gráfico 2, anteriormente apresentado.

Uma vez que o objetivo central deste estudo está focado na identificação de níveis de estabilidade, a seguir, mostra-se que diferentes textos comportam diferentes ecossistemas, ou seja, possuem propriedades que oscilam de acordo com a ação de linguagem dos sujeitos e, para tanto, são apresentados exemplos de três grafos: baixo, médio e alto nível de intersecção e diferença entre o que está no texto-empírico em relação ao conteúdo do texto-referência.

No que se refere ao padrão de estabilidade, investigado por meio da função de Intersecção, na Figura 24, apresenta-se o grafo de intersecção do texto-empírico do T6, o qual evidencia um padrão de estabilidade com percentual inferior ao de outros textos-empíricos, tendo em vista que possui um número pequeno de conceitos, mas, se observado sob outra perspectiva, apresenta Diferença (variabilidade) de um conceito, ou seja, contém um tipo de emergência, mesmo que em proporções menores que, por exemplo, no T10, embora mais extenso não haja.

**Figura 24:** Grafo de intersecção T6



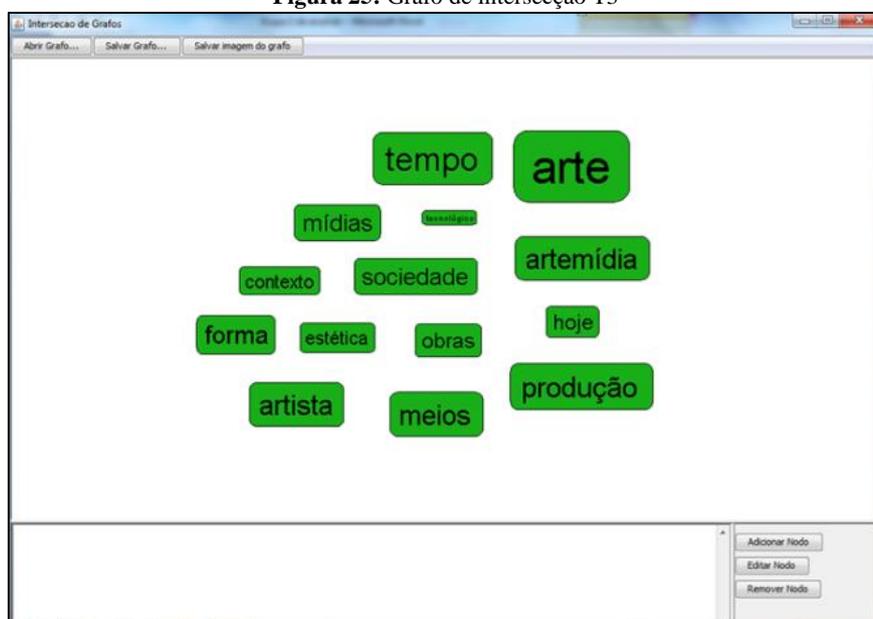
**Fonte:** Autor

Na mesma perspectiva do grafo anterior, a Figura 25, a qual representa o grafo de intersecção do texto-empírico produzido com o T3, deixa evidente um número de conceito superior ao exemplo anterior, sendo possível afirmar que neste caso a ação de linguagem do autor do texto está ainda mais estabilizada em relação ao texto-referência. A partir dessa constatação, pode-se dizer que a emergência sistêmica da interação entre texto-referência e ação de linguagem é discreta, pois afirmar que esta é nula seria o mesmo que dizer que o texto-empírico não é condizente para com a finalidade da

atividade, sendo que, na verdade, a total estabilidade, representa um estado de fase da ação de linguagem do sujeito e não pode ser desconsiderada, pelo contrário, deve ser analisada em relação às zonas desenvolvimentais, propostas por Vygostky (1998).

Neste tocante, apropriar-se dos conceitos e aplicá-los representa preservar e adaptar as condições iniciais das coordenadas do texto-referência às motivações e intenções de elaborar um texto-empírico, que alcance a proposta da atividade de produção textual, que, *a priori*, parece ter como diretriz geral a manifestação dos temas tratados no texto-referência sem esgotá-los, exceto quando o professor propõe esse tipo de abordagem, não sendo este o caso das orientações da atividade de produção textual desta turma. A orientação inicial, a qual não está em questão era ler e produzir um texto autoral.

**Figura 25:** Grafo de intersecção T3



**Fonte:** Autor

A Figura 26, por sua vez, refere-se à produção textual do T10, que deu origem a um grafo com número elevado de conceitos quando comparado aos anteriores, mas que também demonstra que o autor se manteve totalmente estável aos conceitos do texto-referência. Os três exemplos ilustram claramente o padrão identificado na Turma 1 que é o de utilizar o maior número possível de conceitos lidos, com vistas a não fugir do tema e das ideias centrais do texto-referência. No entanto, esse padrão emergiu da análise dos textos-empíricos enquanto uma prática da ação de linguagem de quase todos os sujeitos da Turma 1, os quais manifestaram uma tendência que sugere que permaneceram mais próximos da sua zona de desenvolvimento real do que da zona de

desenvolvimento proximal. Nesse sentido, entende-se que quanto mais variabilidade composicional os sujeitos manifestassem em seus textos, mais emergente teria sido o processo de desenvolvimento e aprendizagem, entendendo-se esses dois últimos conceitos como estados de fase, os quais são transitórios e responsáveis pela consolidação do pensamento humano na atividade de produção textual. Contudo, mesmo estáveis aos conceitos empregados no texto-referência não se pode afirmar que os textos-empíricos sejam lineares, pois as inter-relações e subinter-relações revelaram que a complexidade dos textos está constituída sempre por relações distintas, o que configura uma ação de linguagem autoral e sempre singular (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008).

**Figura 26:** Grafo de intersecção T10



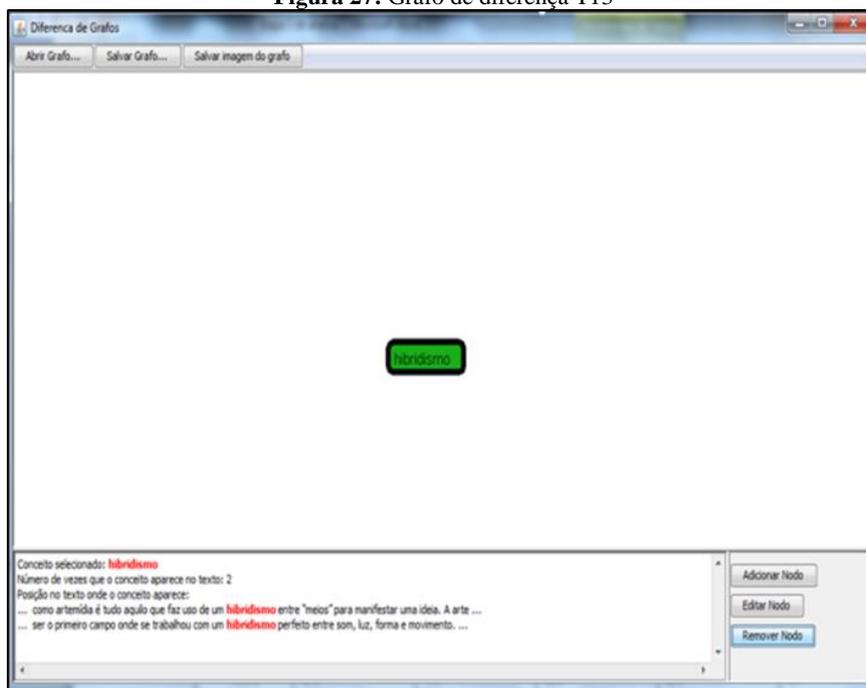
**Fonte:** Autor

As figuras 24, 25 e 26, portanto, ilustram a ideia de que, de fato, a ferramenta Sobek pode isolar e evidenciar, na condição de um grafo, tudo o que estiver presente no texto-empírico, enquanto Intersecção ao conteúdo do texto-referência e seu grafo de recorrências gerais. No que diz respeito à operação entre grafos (e texto-referência), quanto à função de Diferença, essa corrobora a existência de um padrão coletivo de estabilidade do conteúdo dos textos-empíricos, anteriormente afirmado. Conquanto haja a notável emergência de um padrão de estabilidade da ação de linguagem dos sujeitos, por conseguinte, os níveis de complexidade mostraram-se baixos, ou seja, refletem o padrão preponderante. Ressalta-se que, para fins de isolamento dos conceitos de

intersecção e Diferença, a ferramenta, por meio da função *Comparar com texto?*, examina não apenas o que foi tomado como recorrente e apresentado no grafo geral do texto-referência/Tabela 3, mas considera todos os conceitos presentes no texto como estratégia de maior detalhamento quanto à presença de conceitos que não são recorrentes, mas que estão no texto-referência

Em relação ao padrão de variabilidade, investigado por meio da função de Diferença (conceitos emergentes), na Figura 27, apresenta-se o grafo de diferença do texto-empírico T13 que, embora possua um número significativo de conceitos, permitiu observar-se a existência de apenas um conceito (hibridismo) diferente. Essa característica reforça a estabilidade da estrutura dos textos-empíricos, os quais, em geral, apresentam variação entre 1 e 4 conceitos diferentes dos da totalidade de conceitos do texto-referência. Nem mesmo a ação de linguagem dos sujeitos/autores contribui para a ampliação das estratégias enunciativas do texto-empírico (BRONCKART, 2006, 2007, 2008), de modo a torná-los emergentes em um nível distinto da sua condição inicial (texto lido).

**Figura 27:** Grafo de diferença T13

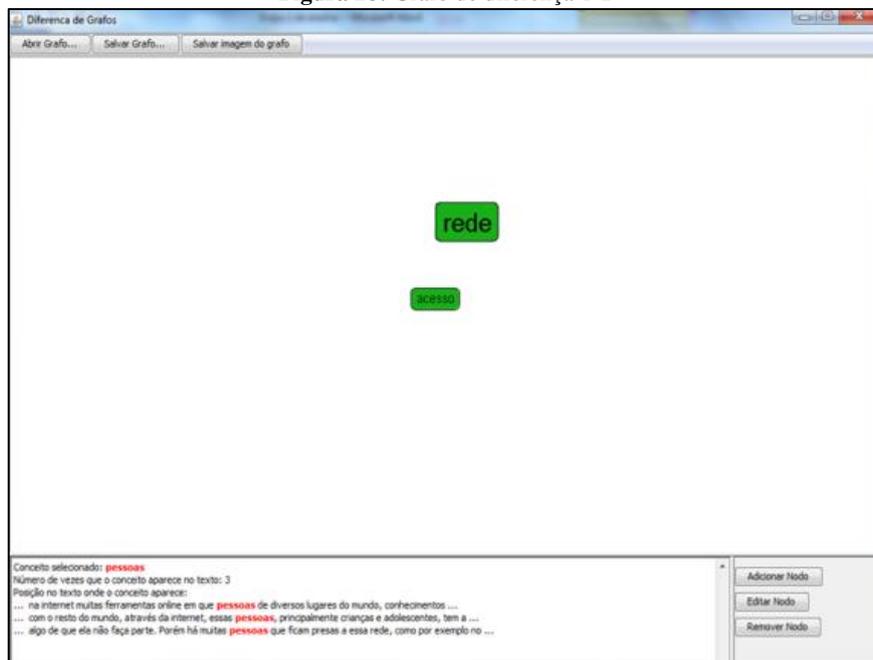


**Fonte:** Autor

Na Figura 28, considera-se que, embora o T2 seja proporcionalmente menor que o T13 (Figura 27), esse apresentou dois conceitos diferentes (rede e acesso) dos do texto-referência. Relativamente a isso, foi possível constatar que muito pouco pode-se inferir acerca do conteúdo de um texto de pouca extensão, posto que há uma série de

interconexões que, se observadas, revelam complexidades até mais significativas do que em textos extensos.

**Figura 28:** Grafo de diferença T 2

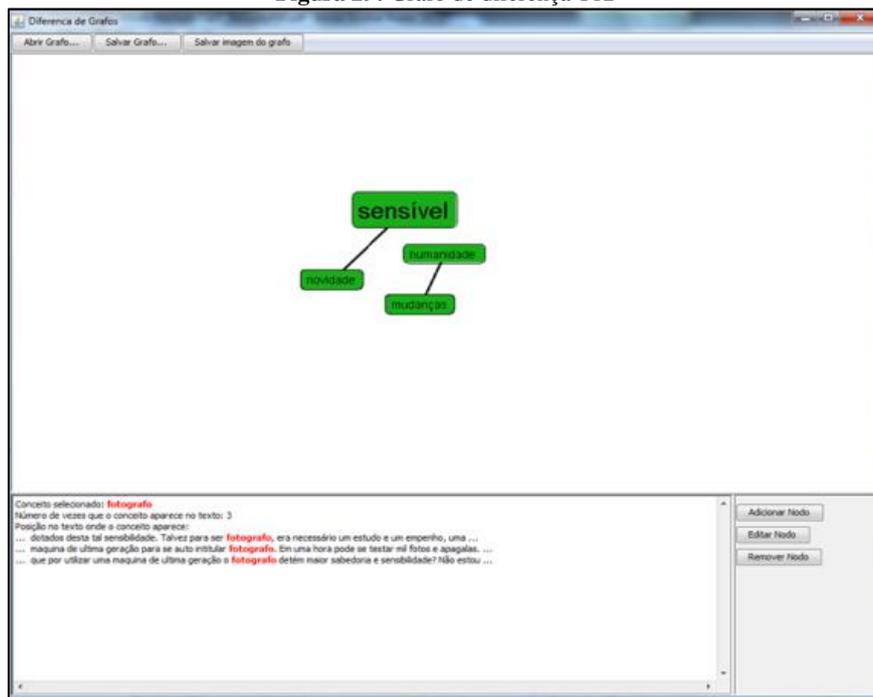


**Fonte:** Autor

Na Figura 29, tem-se o grafo de diferença T12, no qual se identificou a emergência de 4 conceitos (sensível, humanidade, novidade e mudanças), sendo esse o índice mais elevado de conceitos em caráter de variabilidade. Embora ainda constitua um percentual baixo, essa característica, que é relevante para ações de mediação formativa dos sujeitos em contexto formal de educação, torna evidente que há variabilidade na escrita dos textos-empíricos.

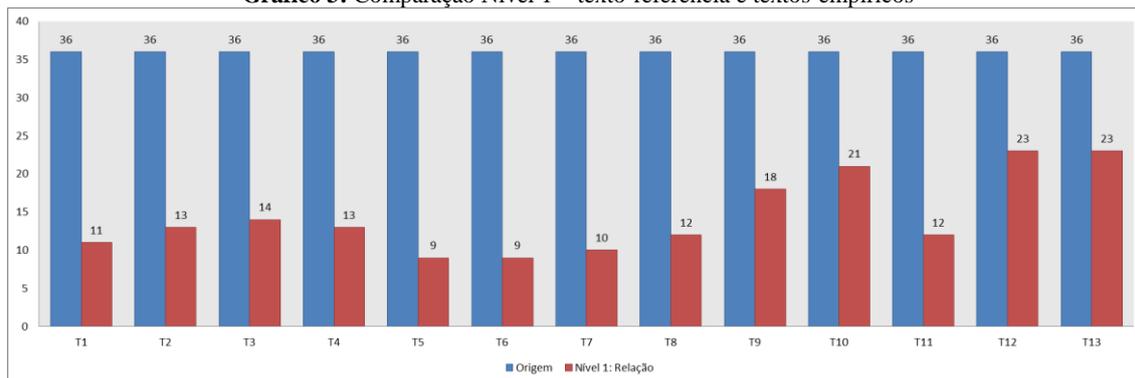
Por outro lado, essa mesma característica remete ao padrão de estabilidade adotado na ação de linguagem de todos os sujeitos da Turma 1, que foi, então, considerado como um dos padrões que emergiram da análise desse grupo. Em suma, a utilização da nova versão da ferramenta, de fato, mostrou ser eficiente na mineração de dados, para fins de observação de complexidades presentes na atividade de produção textual, como atividade de linguagem, bem como para detalhar níveis e padrões presentes nos textos-empíricos (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008). Contudo, é necessário considerar que o processo de mineração demandou refutar conceitos como preposições, conjuntos, números, etc., a fim de dar o devido tratamento aos textos submetidos à referida técnica exploratória.

Figura 29: Grafo de diferença T12



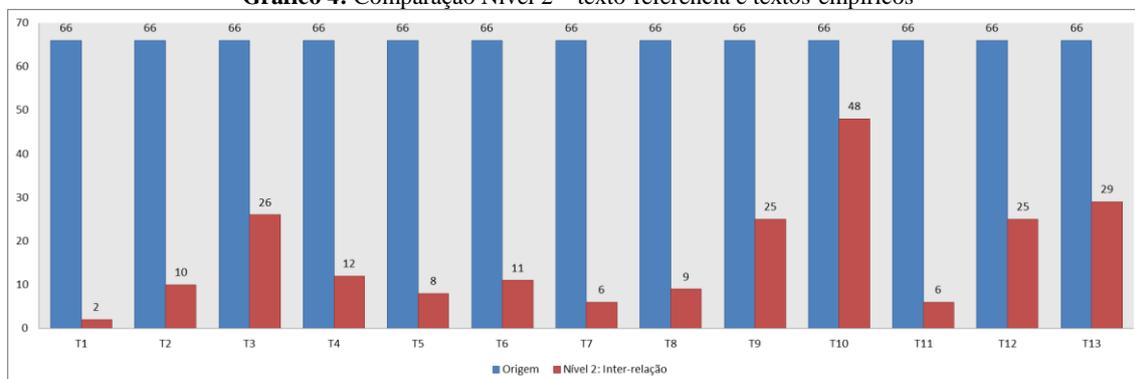
Fonte: Autor

Uma vez que a premissa inicial, que deu origem ao desenvolvimento das funções inseridas nesta versão de Sobek, previa, de modo mais geral, observar a relação do conteúdo do texto-referência em relação ao texto-empírico, considerou-se relevante, apenas para fins de demonstração, exibir gráficos comparativos entre a totalidade de conceitos dos três níveis de complexidade (1, 2, 3) do texto-referência quanto ao número total conceitos de cada nível (1, 2, 3) dos textos-empíricos. Essa estratégia evidencia que a ação de linguagem dos sujeitos mostrou-se quase que completamente distinta, seja em um nível ou em outro. Para tanto, no Gráfico 3, mostra-se a relação do Nível 1 de complexidade do texto-referência, com um total de trinta e seis conceitos, denominando-o como *origem*, uma vez que fora a condição inicial para a produção textual, e o total de conceitos do mesmo nível para cada texto-empírico. Por meio da representação dos índices de cada texto, constatou-se que esses oscilaram entre nove e vinte e três conceitos, os quais, se comparados ao índice do percentual do texto-referência, verifica-se que esse é notadamente menor. Esse aspecto remete a fatores que, embora não possam ser identificados neste estudo, dizem respeito às motivações e intenções que cada autor/sujeito atribuiu ao ato de produzir um texto, seja sendo influenciado pelo conteúdo do texto-referência ou apoiado em suas experiências no mesmo tipo de atividade.

**Gráfico 3:** Comparação Nível 1 – texto-referência e textos-empíricos

**Fonte:** Autor

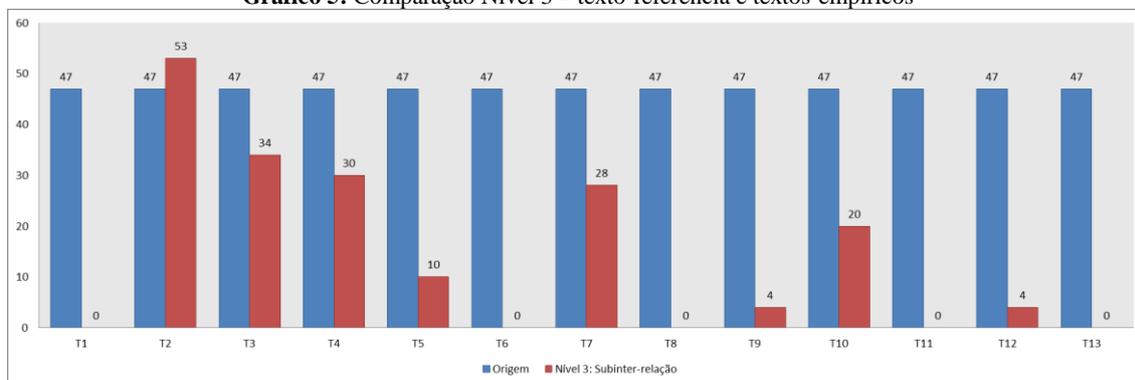
No Gráfico 4, referente ao nível 2 de complexidade, percebeu-se que o índice de inter-relações nos textos-empíricos foi mais baixo que o de conceitos totais do texto-referência, apontando para uma outra instância de complexidade não menos importante que a identificada no nível 1. O Gráfico 3 exibe a relação do Nível 2 de complexidade do texto-referência com um total de sessenta e seis conceitos (origem) e o total de conceitos do mesmo nível para cada texto-empírico. Por meio da representação dos índices de cada texto, constatou-se que eles oscilaram entre dois e quarenta e oito conceitos, ao mesmo tempo em que alguns textos-empíricos apresentaram um padrão de ação de linguagem em menor índice, em outros casos, como nos textos T3, T9, T10, T12 e T13 em que houve significativo aumento desse índice no Nível 2 de complexidade. Assim como no nível anterior, também é possível visualizar-se, graficamente, que sua composição se distingue, não apenas em comparação ao percentual de conceitos do texto-referência, mas também entre si, apontando para a dinamicidade inerente à definição de SACs.

**Gráfico 4:** Comparação Nível 2 – texto-referência e textos-empíricos

**Fonte:** Autor

No Gráfico 5, referente ao nível 3 de complexidade, notou-se que o índice de subinter-relações nos textos-empíricos teve variações mais turbulentas quanto às comparações (Níveis 1 e 2) (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008). Em alguns casos, como nos T1, T6, T8, T11 e T13, não houve a identificação desse nível, mesmo em textos em que, nos níveis anteriores, apresentavam percentuais significativos de conceitos. O Gráfico 5 exibe um total de quarenta e dois conceitos (origem), e o total de conceitos do mesmo nível para cada texto-empírico. Além disso, entende-se que a oscilação entre quatro e cinquenta e três conceitos reafirma a proposição feita neste estudo, quanto à necessidade de aprofundar a perspectiva de mineração de dados com a finalidade de observar outras instâncias que não apenas o grafo inicial gerado pela ferramenta. Também é possível notar neste gráfico que houve em alguns textos-empíricos a manifestação de um índice elevado de conceitos (T2, T3, T4, T7 e T10), mas que, em outros textos, esse nível de complexidade é baixo (T5, T9 e T12) ou inexistente (T1, T6, T8, T11 e T13). De forma semelhante aos níveis anteriores, considera-se que a visualização desses índices é um dado relevante ao usuário da ferramenta/professor, uma vez que propicia uma variável coletiva (padrão), assim como permite um detalhamento da complexidade presente em tais textos.

**Gráfico 5:** Comparação Nível 3 – texto-referência e textos-empíricos

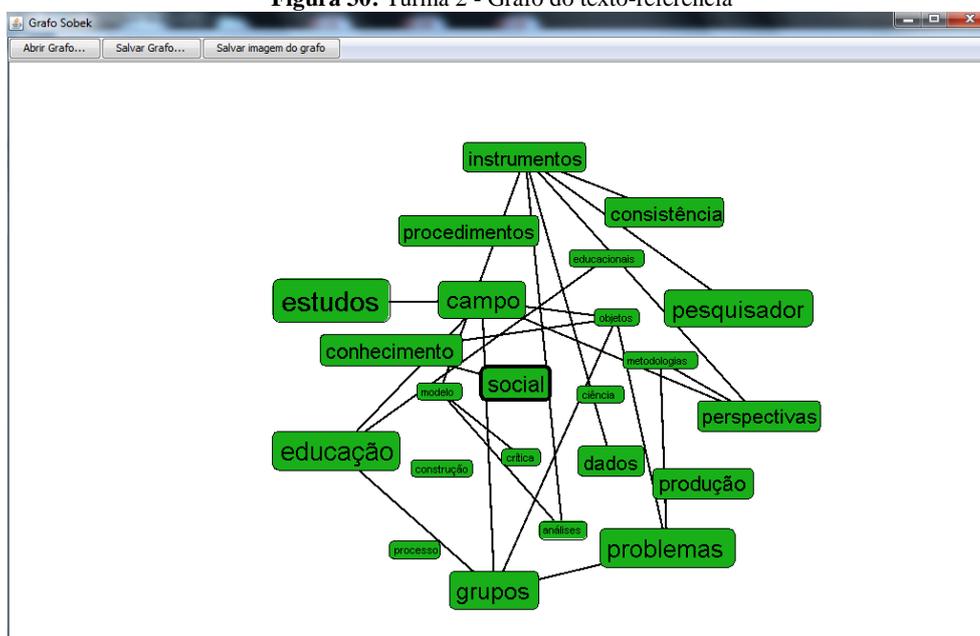


**Fonte:** Autor

Na Turma 2, o texto-referência selecionado pelo professor para a atividade de produção textual foi um artigo científico com a temática sobre Metodologia de Pesquisa em Educação, de curta extensão (8 páginas), o que suscitou um número médio de conceitos (signos) recorrentes em sua estrutura. Seguindo a mesma lógica da atividade de produção textual da Turma 1, após a leitura do texto-referência, também em caráter de atividade extraclasse, foi solicitado aos alunos que produzissem seus textos-empíricos. O total de participantes nessa turma foi de 19 alunos, os quais geraram,

respectivamente, 19 textos-empíricos. Após o recebimento dos textos da T2, primeiramente, o texto-referência (Figura 30) foi submetido ao processo de mineração de dados com a ferramenta Sobek, em sua nova versão, com o objetivo de gerar um grafo composto apenas pelo conteúdo desse texto, sem relacioná-lo a qualquer texto-empírico, com vistas a explorar sua constituição. Em relação ao tratamento dos dados, do conteúdo apresentado nos grafos, assim como referido na abordagem da Turma 1, selecionou-se a opção de extração de cinquenta conceitos por grafo, bem como foram refutados artigos definidos, indefinidos e preposições. Dessa maneira, na Figura 30, apresenta-se um panorama geral dos conceitos recorrentes no texto-referência, compreendidos como pertencentes ao Nível 1 de complexidade, por revelarem a macroestrutura do texto composta pela totalidade de seus conceitos.

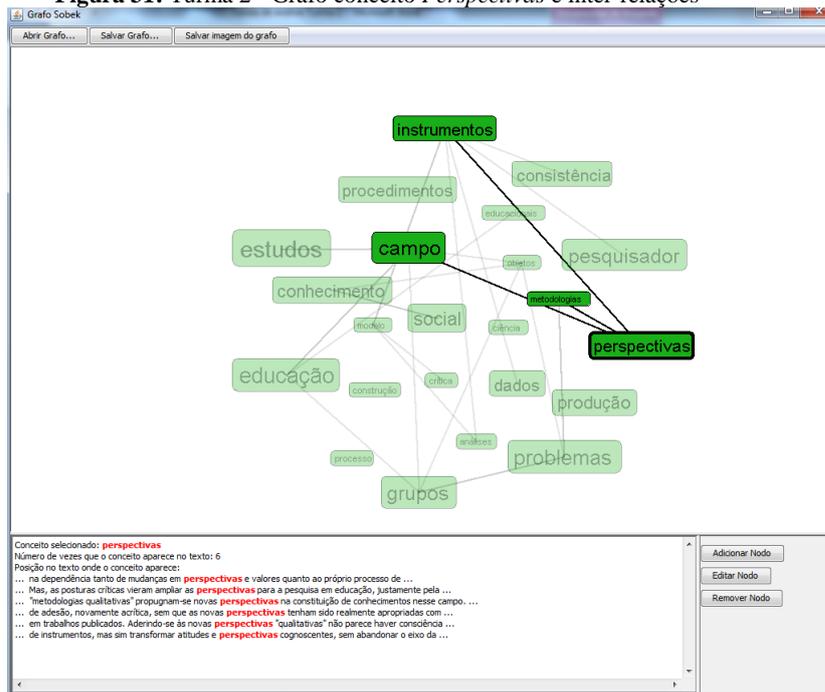
**Figura 30:** Turma 2 - Grafo do texto-referência



Fonte: Autor

Com vistas a analisar o processo de extração e interpretação dos dados do texto referência, em seguida, observou-se a existência do Nível 2 de complexidade, ou seja, a presença de inter-relações. Essa operacionalização torna-se possível ao clicar-se em um conceito (nodo), o qual passa a exibir as inter-relações associadas ao primeiro nível. Para ilustrar essa ação, selecionou-se o conceito *Perspectivas*, inter-relacionado aos conceitos instrumentos, campo e metodológicas (Figura 31).

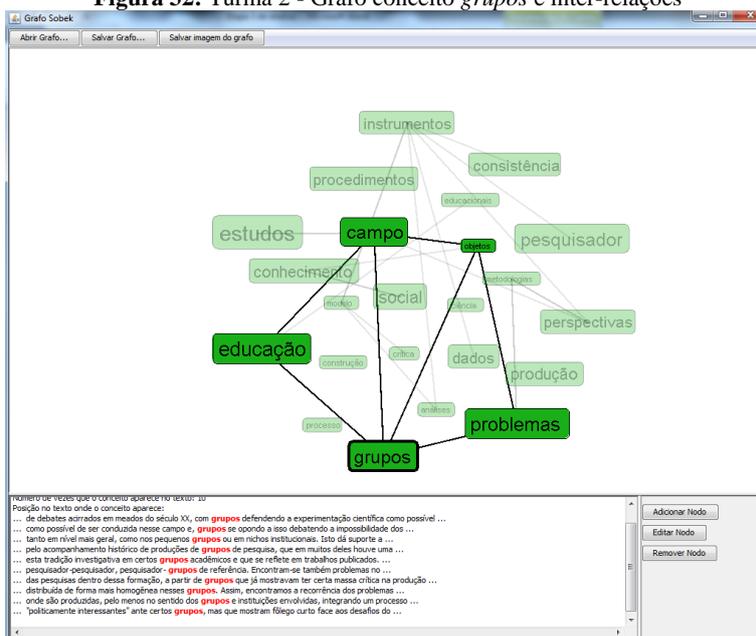
**Figura 31:** Turma 2 - Grafo conceito *Perspectivas* e inter-relações



Fonte: Autor

Com referência ao Nível 3 de complexidade, acerca da visualização de subinter-relações, por meio de associações relacionadas a inter-relações e relações, para ilustrar, selecionou-se o conceito *Grupos*. A partir dele, pode-se analisar uma camada ainda mais fractalizada, que diz respeito à presença de subinter-relações associadas aos níveis 1 e 2. Na Figura 32, podem-se notar quatro subinter-relações (campo, objetos, problemas e educação). Considera-se que os três níveis revelaram os conceitos recorrentes no texto e suas associações, o que remete à ação de linguagem do autor do texto-referência.

**Figura 32:** Turma 2 - Grafo conceito *grupos* e inter-relações



Fonte: Autor

Na tabela 3, assim como proposto na análise do texto-referência da Turma 1, são descritos os referidos três níveis de complexidade, anteriormente, ilustrados por grafos. A descrição (Tabela 4) leva em consideração o conceito enquanto Nível 1 de complexidade, suas inter-relações como Nível 2 e as subinter-relações como Nível 3 do processo de extração de dados.

**Tabela 3:** Níveis de complexidade do Texto-referência (Turma 2)

| <b>Nível 1<br/>(conceito)</b> | <b>Nível 2<br/>(inter-relação)</b>             | <b>Nível 3<br/>(subinter-relação)</b>   |
|-------------------------------|--|---|
| 1. Instrumentos               | Consistência, pesquisador, perspectivas, dados | Modelo, análises                        |
| 2. Procedimentos              | Inexistente                                    | Inexistente                             |
| 3. Consistência               | Instrumentos                                   | Inexistente                             |
| 4. Educacionais               | Educação                                       | Inexistente                             |
| 5. Estudos                    | Campo  | Inexistente                             |
| 6. Campo                      | Estudos  | Objetos, perspectivas, educação, grupos |
| 7. Objetos                    | Inexistente                                    | Campo, conhecimento, grupos, problemas  |
| 8. Pesquisador                | Instrumentos                                   | Inexistente                             |
| 9. Conhecimento               | Objetos, social                                | Inexistente                             |
| 10. Social                    | Conhecimento                                   | Inexistente                             |
| 11. Metodologias              | Perspectivas, problemas                        | Inexistente                             |
| 12. Modelo                    | Inexistente                                    | Instrumento, críticas, análises         |
| 13. Ciência                   | Inexistente                                    | Inexistente                             |
| 14. Perspectivas              | Metodologias, campo, instrumentos              | Inexistente                             |
| 15. Educação                  | Educacionais                                   | Campo, grupos                           |
| 16. Construção                | Inexistente                                    | Inexistente                             |
| 17. Crítica                   | Modelo   | Inexistente                             |
| 18. Dados                     | Instrumentos                                   | Inexistente                             |
| 19. Produção                  | Inexistente                                    | Inexistente                             |
| 20. Processo                  | Inexistente                                    | Inexistente                             |
| 21. Análises                  | Inexistente                                    | Modelo, instrumentos                    |
| 22. Problemas                 | Metodologia                                    | Objetos, grupos                         |
| 23. Grupos                    | Inexistente                                    | Educação, campo, objetos, problemas     |

**Fonte:** Autor

A análise do texto-referência da Turma 2 permitiu identificar um total de vinte e três conceitos, vinte e duas inter-relações e vinte e três subinter-relações. Cada conceito apresenta núcleos distintos (sistemas e subsistemas) de inter-relações e subinter-relações, como visto na Tabela 4, os quais se mostram em acordo com a definição de SACs, adotada neste estudo, quanto à dinamicidade presente na ação de linguagem. Diante da abordagem de utilização da ferramenta de mineração, analisou-se, por meio de grafos, o total de dezenove textos-empíricos produzidos na Turma 2, tendo seus

respectivos níveis de complexidade registrados (GLEICK, 1989; LARSEN-FREEMAN e CAMERON, 2008; MORIN, 2008).

A partir do Gráfico 6, é possível visualizarem-se os percentuais observados em um dos dezenove textos-empíricos minerados, assim como na análise da Turma 1, que também evidenciou a complexidade da ação de linguagem dos sujeitos sob diferentes constituições. Na medida em que alguns apresentam variações discretas, outros se destacam por, *a priori*, já evidenciarem a complexidade processual envolvida na atividade de linguagem realizada. A análise do Nível 1 de complexidade (conceitos) mostrou o engajamento dos sujeitos a partir de motivações e intenções, aparentemente, bem distintas, posto que, nesse nível, o número de conceitos oscilou entre 6 e 21 (T6 e T14), ressaltando-se, porém, que esse aspecto apontou para a constatação de que esse, dentre os demais níveis analisados, a seguir apresentados, foi o de número de conceitos inferior que emergiram da atividade de linguagem proposta.

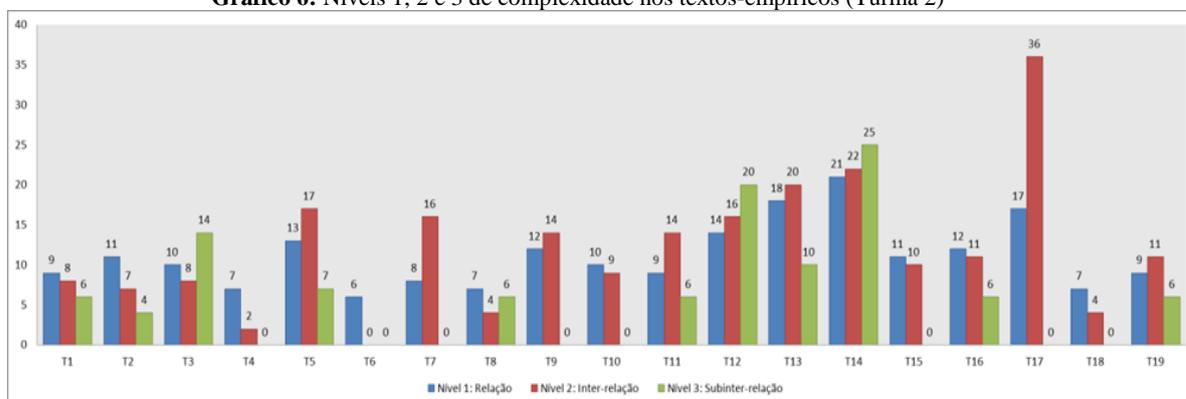
Por outro lado, cabe reiterar que, por ter apresentado um índice inferior de conceitos, este nível não é menos complexo ou menos importante, tanto que os demais estão vinculados a ele. Concernente ao Nível 2 de complexidade (inter-relações), nota-se que há índices que se mantiveram próximos (T1, T2, T3, T8, T10, T15, T16 e T18), afastados (T4, T5, T7, T9, T11, T12, T13, T14, T17 e T19) ou inexistentes (T6) no primeiro nível, tendo variado entre dois e trinta e seis conceitos, porém sobressaindo-se em relação ao nível 3, característica que permite afirmar que, ainda nesse nível, a ação de linguagem se manteve equilibrada, ao passo que, em direção ao último nível, ela revela seu “esgotamento”, ou seja, a finalização do texto.

Quanto ao Nível 3 de complexidade (subinter-relações), notou-se que esse nível oscilou entre seis e vinte e cinco conceitos (T8 e T14), mas, em geral, sua representatividade foi inferior (T1, T2, T5, T8, T11, T13, T16 e T19) ou inexistente (T4, T6, T7, T9, T10, T15, T17 e T18) em relação aos demais níveis, muito embora, em alguns textos, tenham-se identificado índices que se destacam (T3, T12 e T14) quando comparados aos de outros textos. Como dito anteriormente, se se generalizasse esse nível como irrelevante, sob uma perspectiva cartesiana, posto que parece pouco significativo, estar-se-ia desconsiderando parte da complexidade presente não só no processo de mineração de dados, mas no texto como um todo.

A inexistência de subinter-relações em alguns textos remete à hipótese de que nesses a complexidade da ação de linguagem residiu mais nos níveis 1 e 2, uma vez que, ao compor o texto-empírico, o autor necessita estabelecer núcleos temáticos

(conceitos) e o encadeamento discursivo (inter-relações), a fim de que possa narrar, descrever, argumentar acerca de determinado assunto.

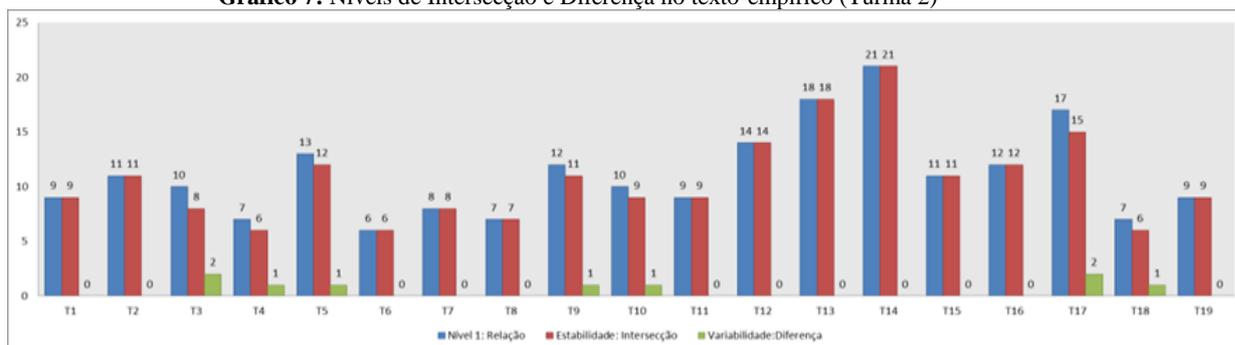
**Gráfico 6:** Níveis 1, 2 e 3 de complexidade nos textos-empíricos (Turma 2)



Fonte: Autor

Assim como na análise da Turma 1, considera-se que o Gráfico 6 corrobora a premissa teórico-epistemológica adotada neste estudo. Isso acontece porque sua representação gráfica da ação de linguagem dos sujeitos permite visualizar concretamente que cada SAC (sujeito ou texto) é singular. Mesmo tendo como base suas experiências coletivas, esses manifestam padrões distintos que, por sua vez, isolados e analisados, podem apontar para a emergência de características coletivas, enquanto variáveis de suas ações, assim como características individuais.

Com vistas a apresentar a análise realizada a partir da inserção das funções de Intersecção e de Diferença em variabilidade, compreendidas como destaque neste estudo, posto que, por meio delas, podem-se apreciar possíveis turbulências/transições/emergências da ação de linguagem dos sujeitos (LARSEN-FREEMAN e CAMERON, 2008), no Gráfico 7, tem-se, primeiramente, o total de conceitos manifestados pelos sujeitos e, em seguida o índice de Intersecção, ou seja, conceitos que se mantiveram em estabilidade ao conteúdo do texto-referência e de Diferença, conceitos que remetem à variabilidade composicional do texto-empírico em relação ao texto lido.

**Gráfico 7:** Níveis de Intersecção e Diferença no texto-empírico (Turma 2)

**Fonte:** Autor

Com a finalidade de aprofundar a perspectiva anteriormente discutida, acerca da dos Níveis 1, 2 e 3 de complexidade dos textos-empíricos produzidos na Turma 2, no Gráfico 7, pode-se notar que os 19 textos-empíricos da Turma 2 são compostos por índices diversos de conceitos, que oscilaram entre seis e vinte e um conceitos, conforme dito na análise do gráfico anterior, na medida em que esses valores são retomados na presente análise, com vistas a examinar a relação do total de conceitos de cada texto-empírico e seus indicativos de estabilidade ou variabilidade conceitual. Uma primeira análise permitiu notar que, de modo geral, assim como na Turma 1, há um padrão de estabilidade na ação de linguagem dos sujeitos, os quais mantiveram os conceitos bastante estáveis em relação ao conteúdo do texto-referência.

Contudo, os índices oscilaram bastante e evidenciaram que, embora haja um padrão de estabilidade na ação de linguagem dos sujeitos (textos), quando submetida ao processo de mineração de dados, acaba por trazer à tona diferentes características em cada nível (1, 2 e 3), como visto no Grafo 6, assim como dentro do padrão de estabilidade (LARSEN-FREEMAN e CAMERON, 2008). Isso porque é possível inferir que o índice de conceitos estáveis oscilou entre seis e vinte e um (T6 e T14), tendo essa característica se sobressaído e se consolidado como um padrão, tecnologicamente observável por meio da função de Intersecção (conceitos que estavam no texto-referência e permaneceram no texto-empírico). Afirmar que há um padrão de estabilidade não significa dizer que não houve apropriação do conteúdo, mas que há uma tendência em manter as propriedades do outro sistema (texto-referência), sem deixar que outras características da ação de linguagem dos autores se façam presentes. No entanto, a ação de deixar-se influenciar mais pelo texto-referência, no sentido de

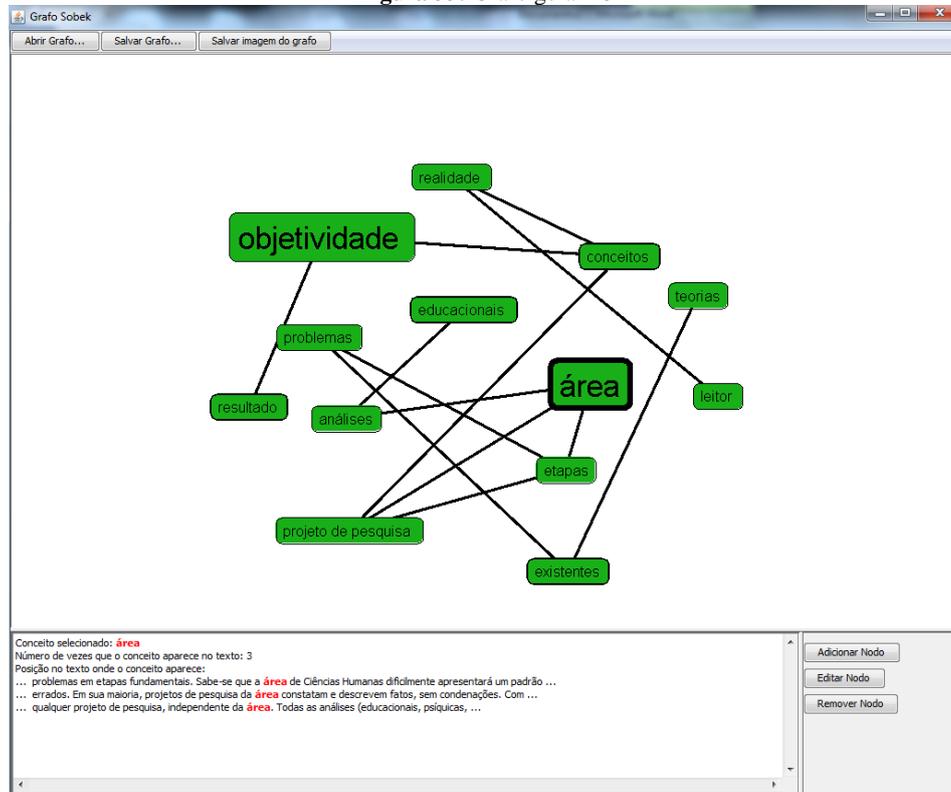
manter seu conteúdo, do que dele se afastar, talvez seja o reflexo das práticas educacionais vivenciadas antes pelos sujeitos.

Tendo em vista que se identificaram níveis de complexidade (1, 2, 3) distintos e um padrão de estabilidade, radicalmente distinto de um padrão de variabilidade conceitual (signos linguísticos), é pertinente ressaltar que, sob a ótica da Teoria da Complexidade, há, na Turma 2, um comportamento complexo, adaptativo e emergente, assim como notado, também, na Turma 1, no sentido de manter-se de acordo com as coordenadas iniciais do texto-referência (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008). A não identificação de um padrão de variabilidade, enquanto conceitos novos são inseridos pela ação de linguagem dos sujeitos, não é, na verdade, a ausência de um comportamento emergente, mas um reflexo da emergência de outro padrão, o de estabilidade. Relativamente ao padrão de estabilidade, é possível notar-se que, na maioria dos textos-empíricos, não houve, na ação de linguagem dos sujeitos, a identificação de conceitos novos (T1, T2, T6, T7, T8, T11, T12, T13, T14, T15, T16 e T19) diferentes dos da utilização na ação de linguagem do autor do texto-referência. Apenas nos textos T3, T4, T5, T9, T10, T17 e T18, notou-se a presença de conceitos distintos da sua condição inicial (texto-referência), mas, ainda assim, em proporções inferiores ao grau de estabilidade. Observa-se, contudo, que a complexidade de um texto não pode ser associada a sua extensão, pois, como visto, textos extremamente pequenos, como os textos T3, T4 e T18, apresentaram a emergência de conceitos novos, ao contrário dos textos T12, T13 e T14, que são de significativa extensão e não apresentaram quaisquer conceitos distintos dos do texto-referência.

Mesmo havendo um padrão de estabilidade tecnologicamente identificável por meio da estratégia de mineração adaptada para os fins desta pesquisa e aplicada aos textos-empíricos, também se notou, nos grafos, uma diversidade composicional distintiva entre os textos minerados e, para tanto, optou-se por apresentar, como feito na análise da Turma 1, grafos específicos dessas questões: a) três grafos que mostram a dinamicidade da ação de linguagem dos sujeitos em perspectiva de índices distintos; b) três grafos que mostram a identificação e a aplicação da função de Intersecção (estabilidade); c) três grafos com a função de Diferença (variabilidade). Com referência à dinamicidade da ação de linguagem percebida nos textos-empíricos no Nível 1 de complexidade (conceitos), verificou-se que o índice de dados oscila. Assim, no T5 (Figura 33), notou-se um índice baixo de conceitos em relação aos demais, ao passo que o T16 (Figura 34) contém um índice médio de conceitos, sem ligações entre eles (os

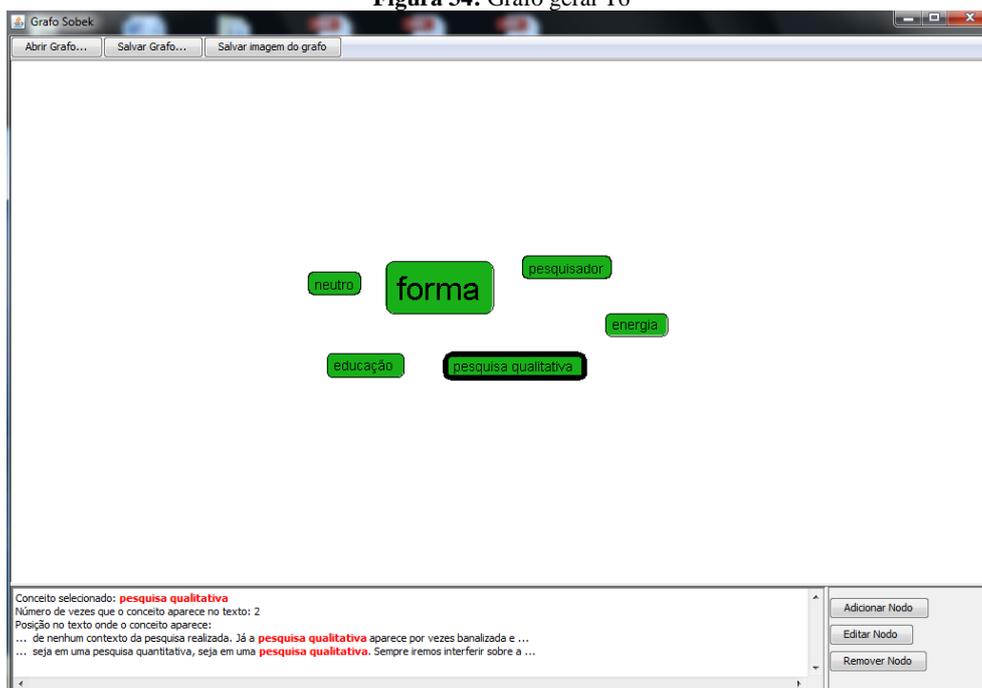
conceitos), devido ao tratamento de dados (refutação de conceitos irrelevantes) que, conseqüentemente, acaba por eliminar possíveis ligações entre conceitos irrelevantes e relevantes. Já no T14 (Figura 35), verifica-se um alto índice de conceitos.

Figura 33: Grafo geral T5



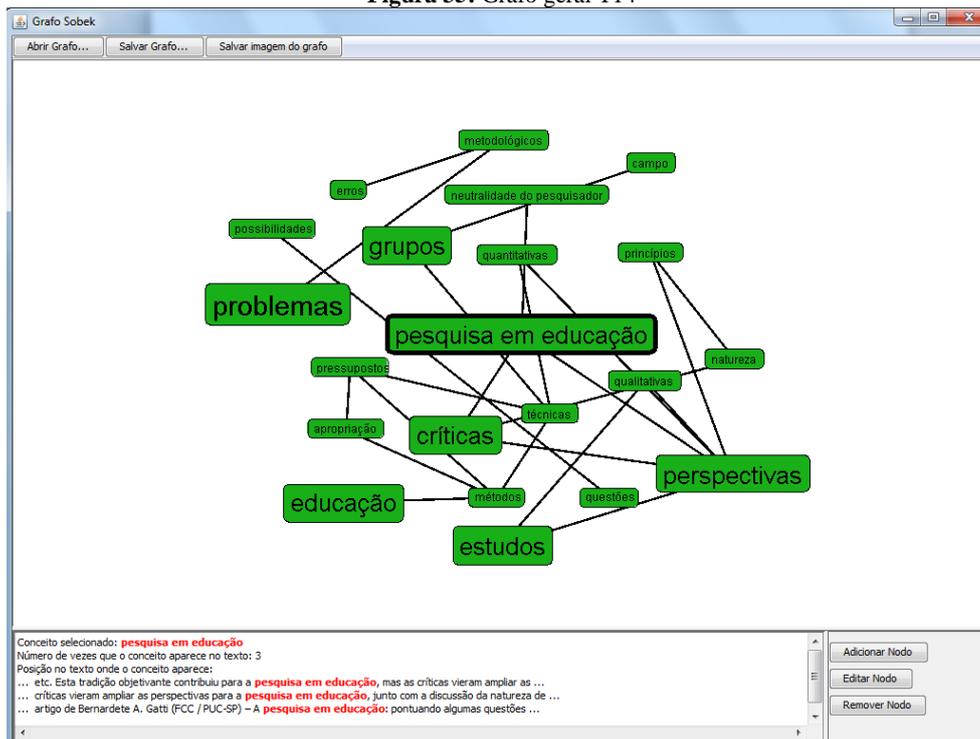
Fonte: Autor

Figura 34: Grafo geral T6



Fonte: Autor

Figura 35: Grafo geral T14

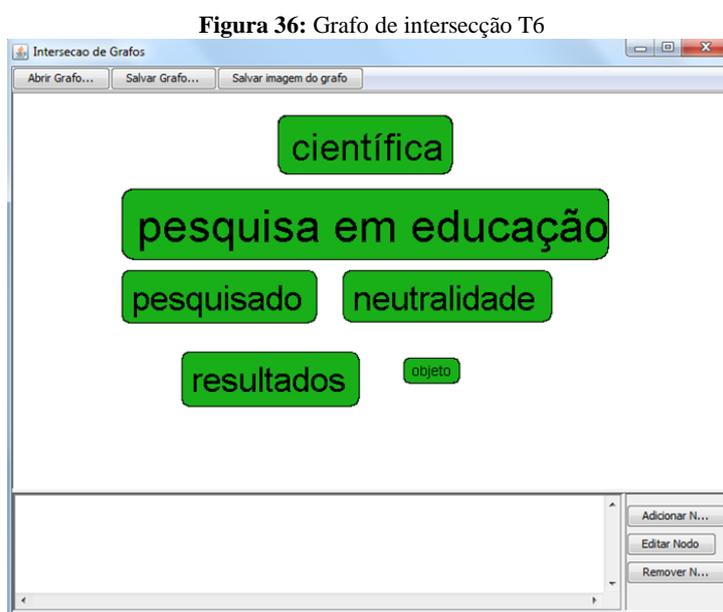


Fonte: Autor

A composição geral dos grafos dos textos T5, T6 e T14 (Figuras 33, 34 e 35) ilustra a ideia de que, embora os sujeitos tenham recebido a mesma condição inicial para a produção textual, as coordenadas da ação de linguagem tomaram formas diferentes e com índices de conceitos distintos, o que reforça a crença de que o processo de apropriação de ideias e o encadeamento delas no texto é sempre uma ação ímpar. Há a presença de trajetórias diversas, pautadas pela ação de linguagem que ressalta, como afirma Vygostky (1998), o fato de que a aprendizagem e o desenvolvimento estão inter-relacionados, uma vez que o sujeito adota pré-construtos (BRONCKART, 2006, 2008), que dizem respeito tanto a uma instância quanto à outra, de modo que essas reverberam nas ações manifestadas nos contextos, produções e ações desempenhadas pelo sujeito. O que se quer dizer com isso é que a possibilidade de organizar o pensamento sob bases distintas não se deve, única e exclusivamente, à condição inicial da atividade proposta, mas também, a fatores múltiplos da constituição/formação prévia de cada sujeito. Os índices de inter-relações e subinter-relações dos textos mencionados podem ser vistos no Gráfico 6, bem como a indicação de níveis de estabilidade ou variabilidade são mostradas no Gráfico 7.

Conquanto este estudo centrou-se na análise de níveis de estabilidade, a seguir, são apresentados exemplos de diferentes grafos obtidos durante a extração de dados

mediada pela ferramenta Sobek. Para fins de ilustração dessa estratégia metodológica, são apresentados três grafos distintos, compostos por níveis diferentes de estabilidade, como referenciado no Gráfico 7. Por meio da utilização da operação de Intersecção entre grafos (e texto), pode-se verificar que, na Figura 36, derivada do texto T6, por exemplo, o nível de conceitos em Intersecção caracterizou-se como inferior aos dos demais grafos (Figuras 37 e 38), apresentados sob a mesma ótica, na medida em que houve total estabilidade à ação de linguagem do autor do texto-referência. Em outras palavras, esse texto-empírico comporta em si apenas a emergência parcial de sua condição inicial, uma vez que não se verificaram conceitos propriamente emergentes entre a tensão do texto-referência e a ação de linguagem de seu autor.

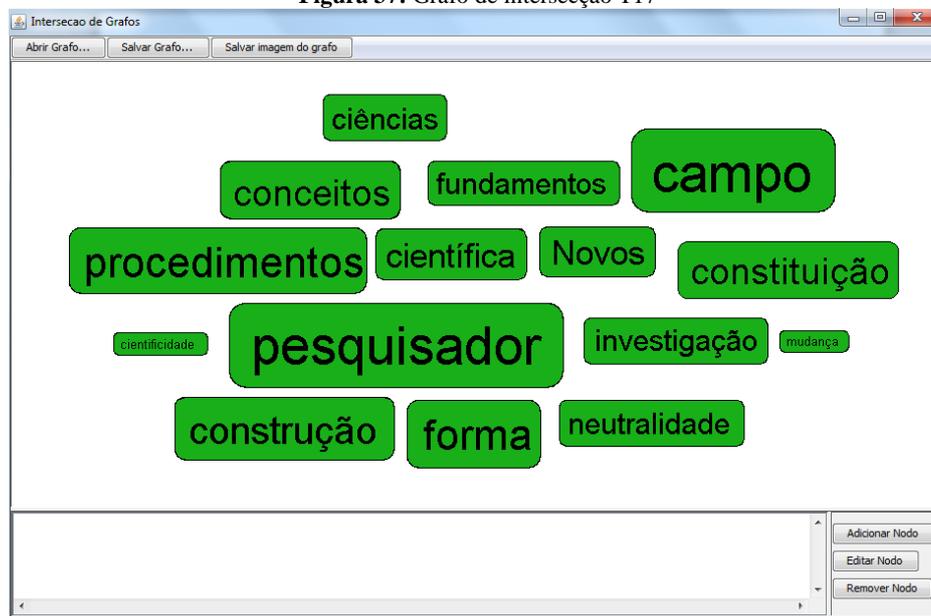


Fonte: Autor

Na Figura 37, que representa o grafo de intersecção do texto-empírico produzido com o T17, visualiza-se que o índice de estabilidade é maior que o apresentada anteriormente, contudo são apresentados conceitos que variaram em relação aos do texto-referência. Nesse tocante, é relevante ressaltar que a configuração da ação de linguagem dos sujeitos em diferentes níveis comporta múltiplas complexidades que, se observadas apenas de um ângulo, podem parecer muito semelhantes, mas elas apresentam dinâmicas composicionais mais caóticas em determinados estados do processo sistêmico. Essa noção está relacionada à abordagem de aproximação dos princípios de Vygostky (1998) à complexidade no que concerne à existência de diversas

zonas desenvolvimentais, mesmo que se tenha de delimitar apenas algumas para operacionalizar essa ou aquela intervenção enquanto mediação formativa.

**Figura 37:** Grafo de intersecção T17

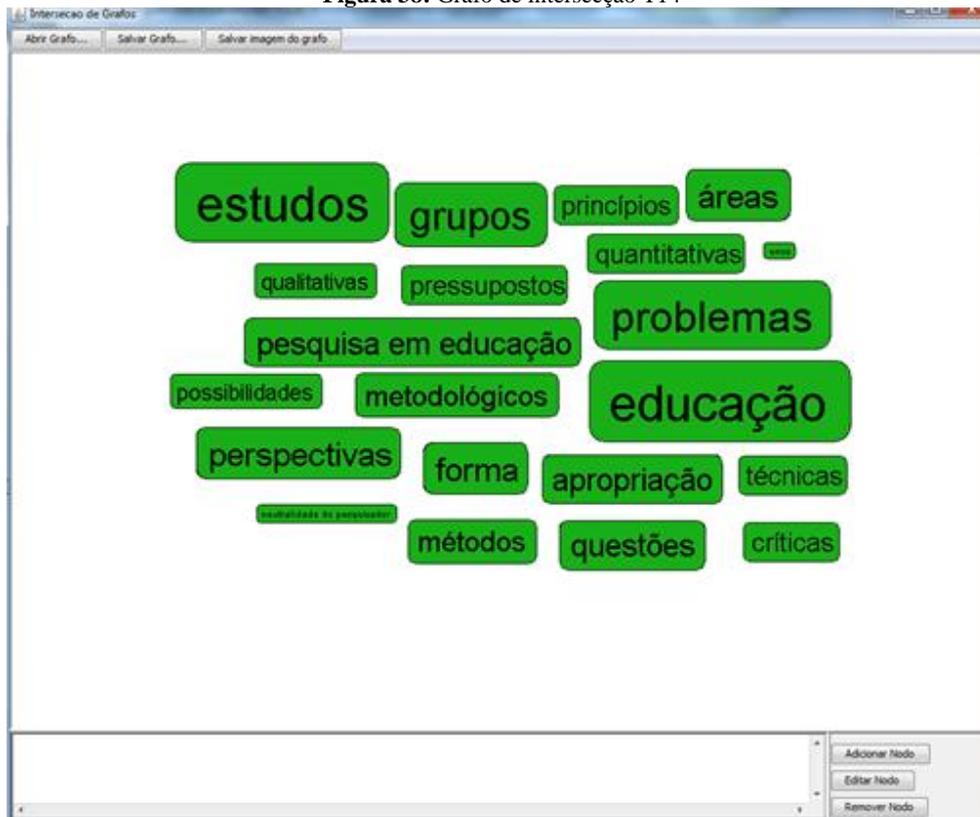


**Fonte:** Autor

A Figura 38 refere-se à produção textual do T14 e contém um índice elevado de conceitos estáveis, aspecto que, mais uma vez, corrobora o padrão coletivo de estabilidade identificado nesta turma. Assim como na Turma 1, na Turma 2, a constatação deste padrão por parte do usuário/tutor/professor pode incitar nele ações pedagógicas que não estão em discussão neste estudo, mas que, se realizadas, podem propiciar benefícios tanto para ações docentes quanto para o desenvolvimento da ação de linguagem dos sujeitos/alunos.

Ações dessa ordem permitem que o professor, em se tratando de contexto educacional, possa prospectar que seus alunos atinjam, de fato, os objetivos de sua disciplina, de modo a conduzi-los à transição dessa Zona Real de Desenvolvimento (ZDR) para uma a Zona de Desenvolvimento Proximal (ZDP), conforme proposto por Vygostky (1998) e discutido ao longo deste estudo. É relevante reafirmar que uma condição/padrão de estabilidade não é algo fixo, mas influenciável por outras coordenadas, compostas por forças entrópicas (textos, alunos, professores, etc.), como postulado pela Teoria da Complexidade.

**Figura 38:** Grafo de intersecção T14



Fonte: Autor

Por meio das figuras 39, 40 e 41 são apresentados exemplos de grafos em que foram percebidas emergências sistêmicas da ação de linguagem dos sujeitos. Se, por um lado, em proporções significativas, tem-se um visível padrão de estabilidade, por outro, notou-se que, em ambas as turmas, houve indícios de que existem emergências, sejam discretas (variabilidade) ou mais acentuadas (estabilidade) (LARSEN-FREEMAN e CAMERON, 2008). Na etapa de exemplificação de grafos em estado de Diferença (variabilidade) da Turma 1, foi possível apresentar grafos compostos com maiores níveis de oscilação (4 conceitos); já na Turma 2, constatou-se que essa realidade oscilou entre um e dois conceitos apenas, sendo, portanto, inferior em proporções de conceitos e de texto com conceitos, em variabilidade.

A aplicação da função de variabilidade, Diferença (conceitos emergentes), ao texto T4, na Figura 39, ilustra apenas um conceito em condição de emergência, mesmo esse sendo um texto de curta extensão, aspecto que, mais uma vez, revela que conceitos distintos aos do texto-referência não estão associados à extensão desse, mas à complexidade da ação de linguagem.

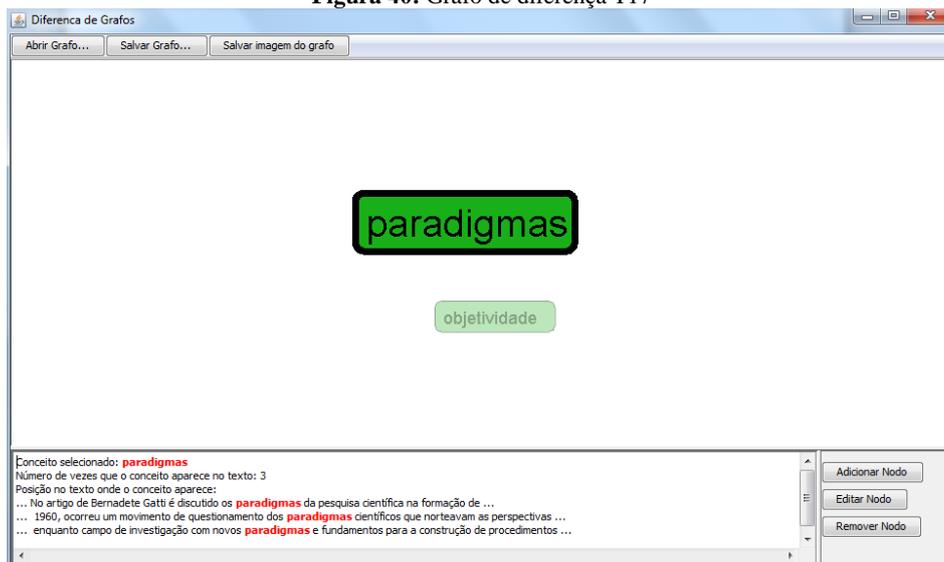
Figura 39: Grafo de diferença T 4



Fonte: Autor

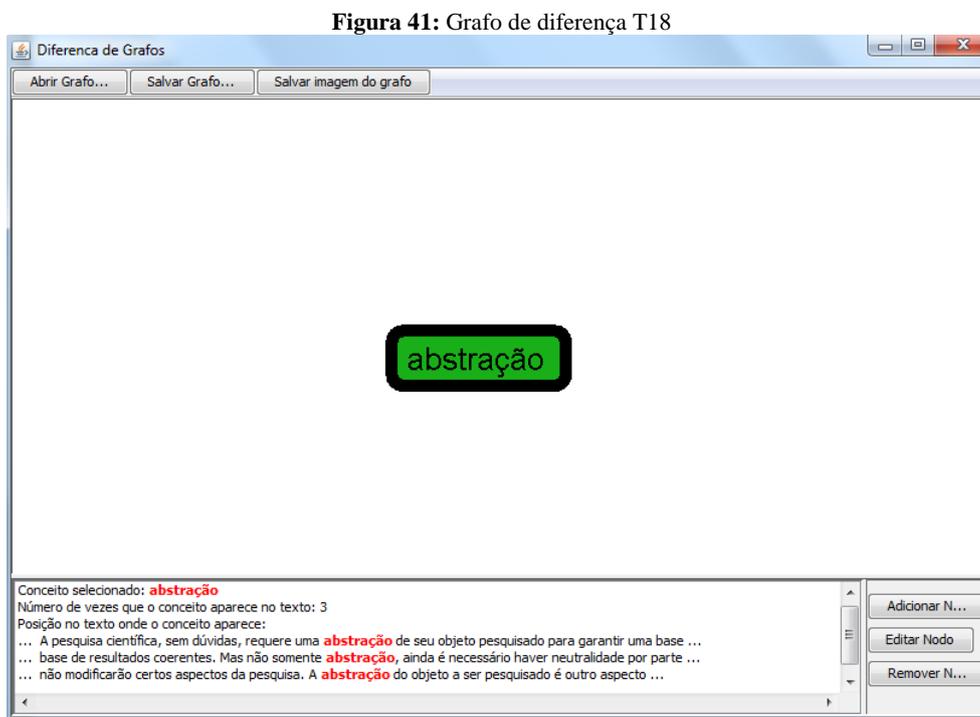
Na Figura 40, derivada da extração de dados do texto T17, obteve-se um total de dois conceitos diferentes dos do grafo e do texto-referência, enquanto índice mais elevado da Turma 2, mesmo sendo um texto composto quase que totalmente por conceitos em estabilidade. A fim de demonstrar que, ao clicar sobre um determinado conceito, pode-se visualizar sua recorrência no texto, selecionou-se o conceito Paradigmas, deixando-o em primeiro plano, acompanhado do conceito *Objetividade*.

Figura 40: Grafo de diferença T17



Fonte: Autor

Na Figura 41, oriunda da mineração de dados do texto T18, verificou-se, também, a existência de apenas um conceito, mesmo sendo esse um texto-empírico com baixo índice total de conceitos em sua estrutura. De modo geral, na Turma 2, a manifestação de conceitos em estabilidade foi considerada inferior à da Turma 1, porém o padrão emergente foi o mesmo.

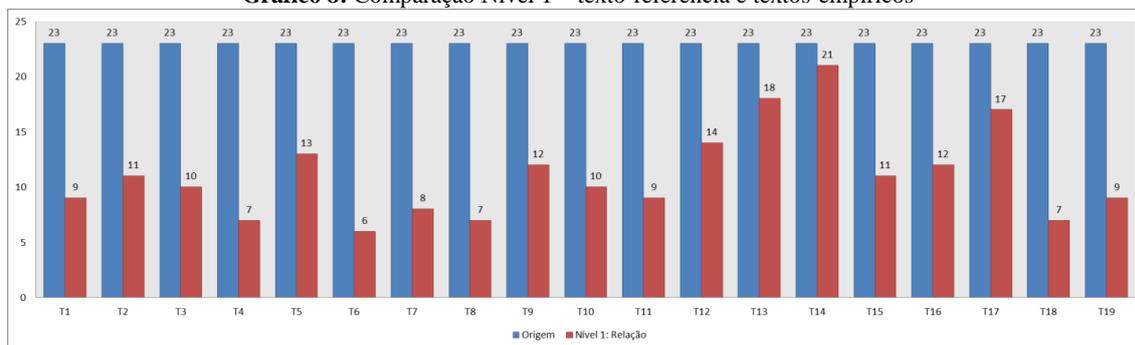


**Fonte:** Autor

Como proposto na análise da turma anterior, para o exame de outra perspectiva investigativa acerca dos níveis de complexidade do texto-referência da Turma 2, quanto a seus índices totais, considerou-se relevante, apenas para fins de demonstração, exibir gráficos comparativos entre a totalidade de conceitos dos 3 níveis de complexidade (1, 2, 3) do texto-referência e quanto ao número total de conceitos de cada nível dos textos-empíricos. Por meio dessa abordagem, foi constatado que a ação de linguagem se configurou como singular em cada texto-empírico, seja no 1, no 2 ou no 3, concepção pela qual se pode visualizar cada texto como um SAC, dotado de características diversas, mas que são resultado da mesma condição inicial, ou seja, da leitura do texto-referência (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008). Portanto, no Gráfico 8, exibe-se a relação do Nível 1 de complexidade do texto-referência, com um total de vinte e três conceitos – denominando-o como *origem*, uma vez que essa foi a condição inicial para a produção textual – e o total de conceitos do mesmo nível para cada texto-empírico. Os índices presentes nos textos variaram entre

seis e vinte e um conceitos, o que deixou visível que alguns textos foram estruturados a partir de índices mais distantes (T1, T4, T6, T7, T8, T11, T18 e T19) ou, aparentemente, mais próximos (T2, T3, T5, T9, T10, T12, T13, T14, T15, T16 e T17) aos do texto-referência. Esse aspecto, na verdade, apenas representa a estrutura geral de construção dos textos-empíricos, sendo esses sistemas diferentes daqueles dos quais extraíram propriedades para dar forma a sua constituição, até porque não se está afirmando que os textos-empíricos deveriam configurar-se como cópias idênticas ou quase idênticas ao texto-referência da atividade de linguagem proposta, nem nesta turma, nem na anterior. Pelo contrário, a apresentação deste gráfico apenas visa a corroborar a concepção de que a tensão gerada pela ação de linguagem quando da atividade de escrita de um texto permite que as características de um processo complexo possam ser isoladas e descritas com base na Teoria da Complexidade e no Interacionismo Sociodiscursivo.

**Gráfico 8:** Comparação Nível 1 – texto-referência e textos-empíricos

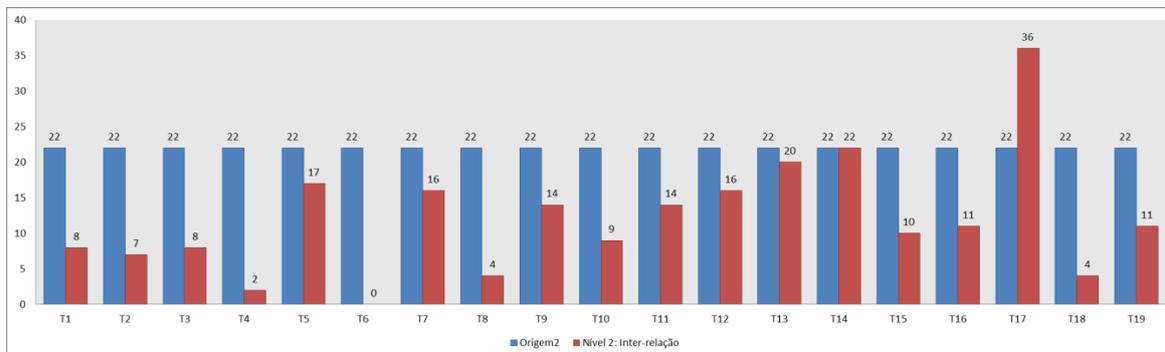


**Fonte:** Autor

No Gráfico 9, referente ao Nível 2 de complexidade, percebeu-se que o índice de inter-relações nos textos-empíricos mostrou-se relativamente inferior ao percentual de conceitos do Nível 1 de complexidade dos textos-empíricos, porém, quando comparados aos do texto-referência, pareceram mais próximos do que no nível anterior, revelando mais uma instância de complexidade da análise realizada. Portanto, o Gráfico 9 exibe a relação do Nível 2 de complexidade do texto-referência com um total de vinte e dois conceitos (origem) e o total de conceitos do mesmo nível para cada texto-empírico. Os índices desses textos oscilaram entre dois e trinta e seis conceitos (T4 e T17). Alguns textos-empíricos apresentaram um padrão de ação de linguagem em menor índice (T1, T2, T3, T4, T8, T10, T15, T16, T18 e T19); em outros casos, como nos textos T5, T7, T9, T11, T12, T13, T14, T17, considerou-se que tiveram índices muito próximos ou superiores aos do texto-referência. A auto-organização e a adaptabilidade da ação de linguagem observada pelo ponto de vista proposto apontam para o fato de que as motivações e intenções dos sujeitos quanto à elaboração do texto-

empírico comportam em si uma gama de complexidades que, talvez, se observadas sem o auxílio da ferramenta Sobek, não seriam notadas.

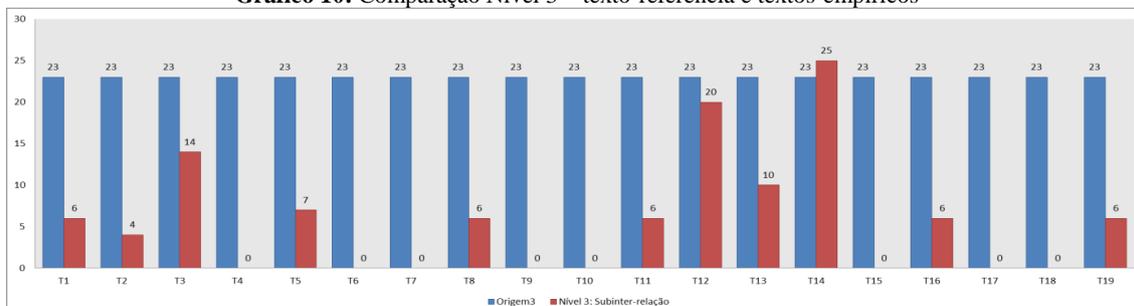
**Gráfico 9:** Comparação Nível 2 – texto-referência e textos-empíricos



**Fonte:** Autor

Em relação ao Gráfico 10 que trata do nível 3 de complexidade, inferiu-se que o índice de subinter-relações nos textos-empíricos absorveu variações mais turbulentas em relação aos níveis anteriores (Níveis 1 e 2) e, até mesmo quando comparado ao nível de conceitos gerais do texto-referência (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008). Em alguns textos, como em T4, T6, T7, T9, T10, T15, T17 e T18, não houve a identificação de conceitos, embora, nos níveis anteriores, parte desses textos tenha exibido índices representativos, aspecto que, novamente, reafirma a condição complexa da ação de linguagem imputada pelos sujeitos. No Gráfico 5, pode-se notar um total de vinte e três conceitos (origem) e o total de conceitos do mesmo nível para cada texto-empírico. Essa análise apresentou índices distintos, porém não tão distintos como os do nível 2, uma vez que se registrou oscilação entre quatro e vinte e cinco conceitos. Vários textos possuíam índices baixos ou médios (T1, T2, T5, T8, T11, T13, T16 e T19) se comparados a outros com números mais elevados de conceitos (T3, T12, T10 e T14). Houve também, em alguns textos (T3, T4, T6, T7, T9, T15, T17 e T18), o registro de total inexistência de conceitos.

**Gráfico 10:** Comparação Nível 3 – texto-referência e textos-empíricos



**Fonte:** Autor

A seguir, apresenta-se a o capítulo de considerações finais.

## 4 CONSIDERAÇÕES FINAIS

A análise dos textos-empíricos, explorados por meio da utilização da ferramenta de mineração de dados Sobek, permitiu corroborar a relevância do aporte teórico-metodológico para a condução dessa tese, que, inicialmente, foi selecionado como abordagem teórica capaz de auxiliar na compreensão da ação de linguagem de sujeitos engajados em atividades de produção de textos, como vistas a demonstrar, concretamente, que um recurso digital pode tornar tecnologicamente observável a complexidade presente em textos, sob a forma de grafos. Diante disso, a premissa unicista de Morin (2008) e Nicolescu (1999) auxiliou no entendimento de que as ciências devem dialogar de modo a promover perspectivas que evidenciem a real complexidade de um determinado objeto de estudo, sem a pretensão de refutar perspectivas provenientes de outros campos do conhecimento, sendo que, neste estudo, afirmou-se que uma ótica complexa e linguística pode, na verdade, ampliar horizontes científicos acerca do modo de interpretar grafos, a partir de um diálogo teórico entre Informática, Educação e Linguística.

O referencial teórico acerca do que se entende por Complexidade e Sistemas Adaptativos Complexos (SACs) permitiu teorizar acerca da ação de linguagem dos sujeitos, sob o viés do Interacionismo Sociodiscursivo (BRONCKART, 2006, 2007, 2008), tornando possível descrevê-la, sob este enquadre, assim como por meio de suas características propor níveis de complexidade, no que se referia à complexidade presente em um grafo de mineração de dados, justamente para tentar expandir a ótica que se pode ter a respeito do conteúdo de um grafo composto por signos linguísticos. A Teoria da Complexidade teve papel central neste estudo, posto que essa perspectiva possibilitou a observação do objeto de estudo, a ação de linguagem, na condição de um estado que, na verdade, não é fixo, mas sempre transitório, aspecto esse que reafirmou a relevância em apresentar a aproximação dos princípios de Vygostky (1998) à essa teoria, pontuando que há múltiplos sistemas, subsistemas, níveis e padrões interconectados e em caráter de recursividade dada a complexidade presente na língua/linguagem.

O exercício de aproximação das afirmações de Vygostky (1998) em relação à Teoria da Complexidade, portanto teve como objetivo mostrar que o desenvolvimento e a aprendizagem são conceitos inter-relacionados e complexos. Contudo, não se objetivou, como afirmado no estudo, identificar e descrever zonas de desenvolvimento,

propriamente ditas, mas sim conjecturar que essas são estados de fase e não, exatamente, SACs. Além disso, inferiu-se que ao apontar níveis de complexidade (1, 2, 3) e padrões complexos (estabilidade e variabilidade) se afirmou que tais zonas podem ser levadas em consideração por quem minera textos, de forma a desenvolver estratégias de mediação formativa para um melhor desempenho (BRONCKART, 2006, 2007, 2008). de quem produz textos, seja no nível individual ou coletivo.

Diante do exposto, o desafio central deste estudo residiu em mostrar que por meio de grafos é possível evidenciar a complexidade (LARSEN-FREEMAN e CAMERON, 2008) presente nos textos-empíricos, e, para tanto, realizou-se no estudo preliminar um exercício de pensar como a referida complexidade poderia ser extraída de modo automatizado por meio da ferramenta Sobek. Então, no estudo preliminar considerou-se necessário desenvolver e inserir funções na ferramenta, a fim de que se pudesse visualizar o conteúdo do texto-empírico em relação ao texto-referência de forma objetiva. Para isso, sob o prisma da própria ferramenta Sobek, postulou-se a existência de três níveis de complexidade (1, 2, 3), como os quais se pode visualizar níveis complexos da ação de linguagem dos sujeitos, demonstrando que o processo de mineração deve levar em conta o conteúdo inicial de um grafo como parte de um todo, porém sem limitar-se apenas a um nível de mineração, mas que conceba que a existência de outros níveis remete a ótica de análise a uma “camada” mais detalhada do processo.

Após esse momento, a utilização das funções inseridas em Sobek permitiu gerar grafos distintos (união, intersecção, diferença), deixando claro ser, plenamente, possível identificar o que no texto-referência configurava-se em caráter de estabilidade ou variabilidade ao conteúdo do texto-referência. Essa abordagem visou isolar o conteúdo do empírico de forma a prover, para quem minera, informações mais precisas acerca da ação de linguagem dos sujeitos. O isolamento do conteúdo estável ou variável do texto-empírico, portanto permite que um professor, por exemplo, possa delinear estratégias mediativas, tendo em mente o quanto o aluno, quando da tensão gerada entre o texto-referência e sua ação de linguagem, manteve sua escrita mais estável ou variável ao conteúdo que lhe serviu de referência. Essa identificação também revela o quanto um padrão de estabilidade ou variabilidade pode revelar acerca do desempenho esperado para uma atividade. Embora não se tenha operacionalizado esse estudo quanto a uma avaliação de desempenho, entendeu-se que a ferramenta pode auxiliar, mais especificamente, docentes em práticas não exatamente avaliativas, do ponto de vista

quantitativo, mas para fins de acompanhamento das trajetórias percorridas pela ação de linguagem individual e coletiva.

A ação de linguagem, materializada nos textos analisados, mostrou-se em um estado caótico (LARSEN-FREEMAN e CAMERON, 2008; BRONCKART, 2006, 2007, 2008), que mesmo mantendo-se estável, em geral, como visto na análise de dados, tende a apresentar, como emergência sistêmica, dados distintos ao texto-referência (atrator caótico), uma vez que as zonas de conhecimentos/desenvolvimento dos sujeitos foram aparentemente influenciadas. Portanto, no que se refere à análise de dados das Turmas 1 e 2, foi possível, com a nova versão de Sobek, notar a presença de um padrão de estabilidade do conteúdo do texto-referência na composição do texto-empírico, muito embora tenha-se mostrado que os textos se distinguiram em extensão e níveis de complexidade (1,2,3), registrou-se o mesmo padrão em ambas as turmas.

Diante desse padrão, considera-se a hipótese de que manter-se mais atrelado ao conteúdo do texto-referência seja uma prática que assegura mais confiabilidade ao aluno, por considerar assim que não estaria desviando-se de seu objetivo discursivo. Por outro lado, infere-se que o ideal não seria um padrão de estabilidade tão significativo, mas sim que o índice de variabilidade fosse relativamente semelhante ou maior, de maneira a inserir novas perspectivas linguísticas e de geração de conhecimento à ação de linguagem, uma vez que segundo Larsen-Freeman e Cameron (2008), na ótica evolucionária da dinâmica da língua os indivíduos enquanto SACs, inseridos em comunidade discursiva, estão entendidos como os principais responsáveis pela noção de mutabilidade e adaptação linguística.

Portanto, gerar grafos distintos de forma automatizada permite verificar a adaptabilidade linguística atribuída pelos sujeitos a sua ação de linguagem, e, com isso por meio dos resultados da mineração pode-se estabelecer com os alunos um diálogo, pós-tarefa, em caráter de mediação. Ainda nesta perspectiva, a identificação viabilizada pela ferramenta possibilita que o professor analise a coerência das aplicações feitas pelos alunos e, não obstante a isso, possa complementar e inserir novos conteúdos alinhados à aprendizagem de alunos em um estágio específico de conhecimento.

A emergência de um tipo de vocabulário propicia que o professor conheça seus alunos e identifique as razões dessa manifestação de complexidade inerente ao agir humano e sua emergência em atividades de produção textual. Por fim, a primeira contribuição desta tese foi a proposição de níveis de complexidade e a utilização de grafos que ressaltam a dinamicidade linguística entre dois textos, e, a segunda, o

desenvolvimento de uma nova versão da ferramenta, a qual tem-se por objetivo disponibilizar na *Internet*. Na continuidade desse estudo, pretende-se propor a utilização desta ferramenta junto a um grupo de professores, como o objetivo de investigar a perspectiva desses e os possíveis resultados empíricos após a utilização de Sobek.

## REFERÊNCIAS

ADAM, J. M., **Eléments de linguistique textuelle**. Liège, Mardaga. 1990

AZEVEDO, B. F. T.; BEHAR, P. A.; REATEGUI, E. B. (2011). **Análise temática das mensagens de discussões online**. In: “VI Congresso Ibero-americano de Telemática”, Gramado. Anais do VI Congresso Ibero-americano de Telemática. Porto Alegre: Instituto de Informática, 2011. v. 6. p. 47-54.

AXT, Margarete; LIMA, José Valdeni; VICARI, Rosa Maria; TAROUCO, Liane Margarida Rosenbach. **Interdisciplinaridade na ótica do Programa de Pós-Graduação em Informática na Educação da UFRGS**. In: PHILIPPI JR, Arlindo; SILVA NETO, Antônio J. Silva. (Editores). *Interdisciplinaridade em Ciência, Tecnologia e Inovação*. Barueri: Manole, 2011.

BALTAR, M. **O conceito de tipos de discurso e sua relação com outros conceitos do ISD**. In: GUIMARÃES, A. M. M.; MACHADO, A. R.; COUTINHO, A. *Interacionismo Sociodiscursivo: questões epistemológicas e metodológicas*. Campinas-SP: Mercado de Letras, 2007.

BAKHTIN, Mikhail Mikhailovitch. **Estética da criação verbal**. São Paulo: Martins Fontes, 2003.

\_\_\_\_\_. **Marxismo e filosofia da linguagem: problemas fundamentais do método sociológico da linguagem**. 12. ed. São Paulo: Hucitec, 2006.

BECKNER, C.; BLYTHE, R.; BYBEE, J.; CHRISTIANSEN, M. H.; CROFT, W.; ELLIS, N. C.; HOLLAND, J.; KE, J.; LARSEN-FREEMAN, D.; SCHOENEMANN, T. **Language is a complex adaptive system: position paper**. *Language Learning*, v.59, supplement 1, 2009. pp.1-26.

BENVENISTE, E. **Problemas de Linguística Geral II**. Campinas: Pontes, 1989.

BERTALANFFY, Ludwing Von. **Teoria geral dos sistemas**. Petrópolis: Vozes, 1973.

BLANCHARD, J. S., FARSTRUP, A. E. **Technologies, Digital Media, and Reading Instruction**. In SAMUELS, S. J., FARSTRUP, A. E. *What Research has to Say about Reading Instruction* (chapter 12). Newark DE: International Reading Association. 2011.

BRONCKART, Jean Paul. **Atividades de linguagem, textos e discursos: por um interacionismo sócio-discursivo**. Trad. Anna Rachel Machado e Péricles Cunha. São Paulo: EDUC, 1999.

\_\_\_\_\_. **Os gêneros de textos e os tipos de discurso como formatos das interações propiciadoras de desenvolvimento**. In: MACHADO, Anna Rachel; MATENCIO,

Maria de Lourdes Meirelles (Org.). **Atividade de linguagem, discurso e desenvolvimento humano**. Campinas: Mercado de Letras, 2006. p. 121-160.

\_\_\_\_\_. **O agir nos discursos**: das concepções teóricas às concepções dos trabalhadores. Tradução de Anna Rachel Machado, Maria de Lourdes Meirelles Matencio. Campinas: Mercado de Letras, 2008.

\_\_\_\_\_. **A atividade de linguagem frente à Língua: homenagem a Ferdinand de Saussure**. In: GUIMARÃES, A. M. M.; MACHADO, A. R.; COUTINHO, A. *Interacionismo Sociodiscursivo*: questões epistemológicas e metodológicas. Campinas-SP: Mercado de Letras, 2007.

CAPRA, Fritjof. **As conexões ocultas**. São Paulo: Cultrix, 2002.

CORRÊA, Ygor. **O agir linguageiro na perspectiva dos sistemas adaptativos complexos em ambiente virtual de aprendizagem em EAD**. Dissertação (mestrado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Linguística Aplicada, São Leopoldo, RS, 2012.

CRISTOVÃO, V. L. L. **Procedimentos de análise e interpretação em textos de avaliação**. In: GUIMARÃES, A. M. M.; MACHADO, A. R.; COUTINHO, A. (Org.). *O interacionismo sociodiscursivo*: questões epistemológicas e metodológicas. Campinas: Mercado de Letras, 2007. p. 257-272.

CRESWELL, John W. **Projeto de Pesquisa**: Métodos Qualitativo, Quantitativo e Misto. 2ª ed., Porto Alegre: Artmed, 2007.

FELDMAN, R., SANGER, J. **Text Mining Handbook**. Cambridge, Inglaterra: Cambridge University Press. 2006.

FRAGA, Dinorá Moraes de. **Um estudo sociointeracionista da emoção, na emergência da internet como contexto de produção textual**. Revista da ABRALIN, Brasília, DF, v. 6, n. 2, p. 215-233, jul./dez. 2007a.

\_\_\_\_\_. Inserção epistemológica da teoria da complexidade nos estudos linguísticos: sobre fragmentos e totalidades. **Calidoscópico**, São Leopoldo, v. 5, n. 3, p. 185-189, set./dez. 2007b.

\_\_\_\_\_. A internet como contexto de produção textual: possíveis implicações do ISD. **Calidoscópico**, São Leopoldo, v. 2, n. 2, p. 55-59, jul./dez. 2004.

GLEICK, James. **Caos**: a criação de uma nova ciência. 16. ed. Rio de Janeiro: Campus, 1989.

GOLDENBERG, Mirian. **A arte de pesquisar**: como fazer pesquisa qualitativa em Ciências Sociais. 2.ed. Rio de Janeiro: Record, 2001.

HABERMAS, J. **Théorie de l'agir communicationnel**, t. I e II. Paris, Fayard. 1987.

HESSLER, J.; REATEGUI, E. **Um método para o apoio à leitura baseado no uso de uma ferramenta de mineração de texto.** In: “Novas Tecnologias na Educação”; CINTED, V. 8 N° 3, dezembro, 2010.

KLEMANN, Miriam; LORENZATTI, Alexandre; REATEGUI, Eliseo. **O Emprego da Ferramenta de Mineração de Textos SOBEK como Apoio à Produção Textual.** In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 20º, Florianópolis, 2009. Anais. Disponível em: <<http://www.br-ie.org/pub/index.php/sbie/article/view/1154/1057>>. Acesso em: junho. 2015.

KLEMANN, Miriam; REATEGUI, Eliseo; RAPKIEWICZ, Clevis; **Análise de Ferramentas de Mineração de Textos para Apoio à Produção Textual.** In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 22º, Aracaju. 2011. Anais. Disponível em: [www.brie.org/pub/index.php/sbie/article/download/1866/1632](http://www.brie.org/pub/index.php/sbie/article/download/1866/1632)>. Acesso em: setembro 2015.

KOCH, I. G. V. ELIAS, V. M. **Ler e compreender os sentidos do texto.** 3.ed. São Paulo: Contexto, 2010.

LARSEN-FREEMAN, Diane. Chaos/complexity Science and second language acquisition. **Applied Linguistics**, Oxford, v. 18, n. 2, p. 141-165, June 1997.

LARSEN-FREEMAN, Diane; CAMERON, Lynne. **Complex Systems and Applied Linguistics.** Oxford: Oxford University Press, 2008.

LEFFA, Wilson J. Se muda o mundo muda: ensino de línguas sob a perspectiva do emergentismo. **Calidoscópico**, São Leopoldo, v. 7, n. 1, p. 24-29, 2009.

LÉVY, Pierre. **Cibercultura.** São Paulo: Ed. 34, 1999.

LORENZ, E. N. **The essence of chaos.** Seattle: Univ. Washington Press, 1993

MACEDO, A. L.; REATEGUI, E. B.; LORENZATTI, A.; BEHAR, P. A. **Using Text-Mining to Support the Evaluation of Texts Produced Collaboratively.** In: Arthur Tatnall; Anthony Jones. (Org.). “Education and Technology for a better world”. 1 ed. Berlin /Germany: Springer, v. 1, 2009. p. 368-377.

MACEDO, A. L.; GRACIOLI, F.; REATEGUI, E. B.; BEHAR, P. A.. **Estratégias pedagógicas de apoio à leitura, à escrita e ao acompanhamento baseadas na tecnologia de mineração de texto.** In: Anais do XXII SBIE - XVII WIE, nov. 2011

MANDELBROT, B. B. **The fractal geometry of nature.** New York: W.H. Freeman and Company, 1982.

MARCUSCHI, Luiz Antônio. **Produção textual, análise de gêneros e compreensão.** São Paulo: Parábola Editorial, 2008.

MARTINS, A. C. **A emergência de eventos complexos em aulas on-line e face a face: uma abordagem ecológica.** Tese (Doutorado) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte. 2008

MORIN, Edgar. **Ciência com consciência**. Rio de Janeiro: Bertrand Brasil, 1996.

\_\_\_\_\_. **Introdução ao pensamento complexo**. 5. ed. Lisboa: Instituto Piaget, 2008.

MORIN, Edgar; LE. MOIGNE, Jean Louis. **A inteligência da complexidade**. 2. ed. São Paulo: Petrópolis, 2000.

\_\_\_\_\_; \_\_\_\_\_. **Inteligência da complexidade: epistemologia e pragmática**. Lisboa: Instituto Piaget, 2009.

NICK. C. Ellis; LARSEN-FREEMAN, Diane (Ed.). **Language as a complex adaptive system**. Malden, MA: Wiley-Blackwell, 2009

NICOLESCU, B. **Uma visão do mundo: a transdisciplinaridade**. IN: O Manifesto da Transdisciplinaridade. Trad. Lucia Pereira Souza. São Paulo: TRIOM, 1999. Disponível em < <http://ciret-transdisciplinarity.org/transdisciplinarity.php> > Acesso em jan.2015

PAIVA, V. L. M. de O. **Autonomia e complexidade**. Linguagem e Ensino, v. 9, n.1, p. 77-127, 2006.

\_\_\_\_\_. **Caleidoscópio: fractais de uma oficina de ensino aprendizagem: memorial**. 2002. Belo Horizonte: UFMG/FALE, 2002. 262 f. Memorial apresentado para concurso de Professor Titular na Faculdade de Letras da Universidade Federal de Minas Gerais.

\_\_\_\_\_. 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.

PARREIRAS, V. **A sala de aula digital sob a perspectiva dos sistemas complexos: uma abordagem qualitativa**. 2005. 356 f. Tese (Doutorado em Linguística Aplicada) – Faculdade de Letras, UFMG, Belo Horizonte, 2005.

PRIGOGINE, Ilya, STENGERS, Isabelle. **A nova aliança**. Brasília: Editora UNB, 1997

REATEGUI, E., PINHO, I. EPSTEIN, D. **Sobek: a text mining tool for educational applications**. In: INTERNATIONAL CONFERENCE ON DATA MINING, Las Vegas, Nevada, USA. Anais. Las Vegas: [s.n.], 2011. p. 59-64.

RICOEUR. **Le discours de l'action**, in Ricouer, P. (ed), La sémantique de l'action. Paris, cnrs. 1977

SADE, Liliane. A. R. **Comunidades, identidade e aprendizagem de inglês sob a ótica do caos e dos sistemas complexos**. 2009. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2009.

SCHENKER, A. **Graph-Theoretic Techniques for Web Content Mining**. PhD thesis, University of South Florida. 2003

SCHNEUWLY, Bernard; DOLZ, Joaquim. **Gêneros orais e escritos na escola**. Campinas: Mercado de Letras, 2004.

SILVA, Valdir. **O conceito de Zona de Desenvolvimento Proximal (ZDP) na perspectiva da teoria da complexidade e do caos: uma releitura** IN: Paiva, V. L. M. O; Nascimento, M. (Org.) **Sistemas adaptativos complexos: língua(gem) e aprendizagem**. Belo Horizonte. Faculdade de Letras da UFMG, 2009

\_\_\_\_\_. **A dinâmica caleidoscópica do processo de aprendizagem colaborativa no contexto virtual: Um estudo na perspectiva da complexidade/caos**. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

SILVERMAN, David. **Doing qualitative research: a practical handbook**. London: SAGE, 2000.

SOUZA, Valeska. **Dinamicidade e adaptabilidade em comunidades virtuais de aprendizagem: uma textografia à luz do paradigma da complexidade**. Tese (Doutorado em Estudos Linguísticos)– Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2011

SAUSSURE, F. **Cours de linguistique générale**. Paris, Payot. 1916

TAN, A. H. **Text Mining: The state of art and the challenge**. In: Proceedings of the PAKDD, Workshop of Knowledge Discovery. 1999

TORRANCE, M.; GALBRAITH, D. **The processing demands of writing**. In: MACARTHUR, C.; GRAHAM, S.; FITZGERALD, J. (Ed.) **Handbook of writing research**. New York: The Guilford Press. 2006

VETROMILLE-CASTRO, Rafael. **A interação social e o benefício recíproco como elementos constituintes de um sistema complexo em comunidades virtuais de aprendizagem para professores de língua**. 2007. 226 f. Tese (Doutorado em Informática na Educação) – Faculdade de Educação, Programa de Pós-Graduação em Informática na Educação, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2007.

VYGOTSKY, L.S. **A formação social da mente**. São Paulo: Martins Fontes, 1998

WALDROP, M. M. **Complexity: the emerging science at the edge of order and chaos**. New York: Touchstone, 1992 Disponível em: < <http://www.finchpark.com>>. Acesso em: 05 fev. 2015.