Elizabete Damasceno¹

Semântica, lexicometria e representação do conhecimento: congruências e aplicações ao ensino de língua materna

Resumo

Os instrumentos de tratamento informatizado de dados textuais abrem novas perspectivas para os métodos de análises de dados qualitativos. Este documento discute as operações para a reorganização formal das sequências textuais e das análises estatísticas aliadas à análise de conteúdo, lexicometria e à Semântica para estudar o vocabulário de um corpus de textos, com aplicações ao ensino de línguas.

Abstract

The statistics and textual data analysis open new opportunities for content analysis. This document discuss the operation to formal reorganization of the texts sequences and of the statistics analysis allied textual data analysis, lexicometry and Semantic for studying the vocabulary of the textbody with application to language education.

Palavras-chave: lexicometria; análise de conteúdo; semântica e ensino de línguas.

Keywords: *lexicometry; content analysis; semantic and language education.*

1. Introdução

Sabe-se que o sistema linguístico é um dos sistemas de significação humana. Outros sistemas de significar o mundo, o homem e o conjunto de relações advindas existem e estão em constante atuação nas práticas humanas historicamente construídas. Nesse aspecto, criar a si mesmo histórica e socialmente como um ser dotado de sentido e de potencialidades, parece-nos tarefa de um sujeito em permanente construção e reconstrução de seu saber e de seu saber-fazer, que, no complexo social, linguístico e cultural, são ações dinamizadas tanto por recortes e relações conceituais para a produção de novas significações, quanto pelas articulações de diferentes redes significativas em circulação no cenário social, em semiose ilimitada.

Em sociedade, a dinamicidade dos diferentes e múltiplos agrupamentos sociais atesta a existência de uma variabilidade de discursos-concretos, produzidos e reproduzidos no quadro geral da comunicação humana, que, pela correlação espaço <> tempo, possibilita a apreensão de diferentes e múltiplas significações, visão de mundo e recortes culturais. Essa relação entre espaço <> tempo motiva, constantemente, o aparecimento e a troca de um número considerável de construções léxico-sintático-semânticas propostas por um duplo significante: sujeito de um fazer e também espectador desse fazer.

¹ Elizabete Damasceno - Professora Doutora de Língua Portuguesa, Prof. Pesquisador da FICS, onde desde o primeiro semestre de 2012 compõe a grupo de pesquisa – oficina de artigos e Iniciação Científica das Faculdades Integradas Campos Salles – FICS. www.campossalles.edu.br

Duas ações - individual e coletiva – possibilitam a existência de uma sintaxe das significações humanas, em suas relações interiores e exteriores, de que resultam a integração e a interação dos sujeitos aos espaços que lhe são correspondentes, pela produção de mensagens plurissignificativas e, por isso, plausíveis. Eis operações construtoras de rede de leitura constituída de saberes culturais, relações sensíveis e julgamentos com os quais potencialidade de significação e enunciados (re)encontram-se e são ajustados por um “eu”, que coloca em movimento não só projetos de ação e de interação e visadas de saber, como também um conjunto de emoções, sob a ordenação de uma configuração complexa: a figurativização de objeto(s) de valor para a construção do tecido discursivo.

Nesse cenário, constroem-se e consolidam-se movimentos de participação, responsabilidade e compromisso social, a fim de que se possa atender à diversidade e às diferenças, aos desejos, necessidades e demandas da modernidade e da tradição, em que conteúdos “conceptuais”, “factuais”, “procedimentais” e “atitudinais” constituem reflexo e seleção de aspectos da cultura, e contribuem para o desenvolvimento da individualização e socialização de competências e habilidades em diversas e diferentes “áreas” e “domínios de saber”, que não podem se apresentar descontextualizadas da construção e do desenvolvimento cultural, para mais além da esfera individual ou pessoal, no âmbito de finalidades e tarefas compartilhadas, compromissos e implicações, considerados acordos e contratos consensuais, a exemplo das línguas naturais.

2. Semântica e representação do conhecimento

Dos diferentes discursos manifestos em diferentes códigos produzidos e reproduzidos socialmente, muitas significações são também produzidas e, recuperadas pelos sistemas

semióticos, armazenadas como grandezas-signos na memória do(s) falantes(s) convertidas em modelos disponíveis para atualização e em modelos para nova(s) análises(s) dos dados da experiência, caracterizado(s) por constantes movimentos de reflexão, projeção, recepção e ação.

Dentre outros, o código linguístico, considerada a sua natureza e o seu funcionamento, apresenta-se como um sistema de significação altamente produtivo: elabora, constrói, estrutura e reformula, permanentemente, os dados da experiência, considerados os diferentes *topoi*, cujos resultados de recorte(s) sobre substância amorfa são novos recortes conceptuais, tanto em relação ao processo de estruturação de novas funções semióticas e metasemióticas (PAIS: 1993), como em relação à codificação de novas informações, que, como já se afirmou, são recuperadas pelo sistema e armazenadas no banco de memória do(s) falantes para suas novas atualizações, constantes e ilimitadas.

Pela produção de atos comunicativos, interativos, reconhece-se o dinamismo das estruturas linguísticas, ou dito de outro modo, reconhecem-se a estruturação e a reestruturação de unidades léxicas, definidas como unidades memorizadas, em competência, pelo falante, uma exigência do próprio meio sócio-cultural, em constante evolução.

Entendemos que entre Universo Léxico de uma comunidade considerada e estruturas sócio-culturais há uma relação de interdependência, que pode ser verificada, compreendendo-se o conjunto dos elementos efetivos e virtuais do sistema linguístico considerado; a combinatória e as exigências da distribuição semiotática; o processo e condição da produção discursiva e seu próprio funcionamento no espaço da vida social; a norma e registros linguísticos empregados em relação também ao contexto e à situação vivenciados pelo(s) falante(s) em seus atos de

comunicação, considerados os vários tipos e gêneros textuais em circulação na sociedade.

No espaço discursivo, vertentes figurativas possibilitam o aparecimento do sentido como conteúdo de um estado inicial a transformar-se e de uma estado final transformado e contextualizado, fundantes de significação e de verificação histórica, autorizados e sustentados que são por uma competência cultural. Eis rede de linguagem complexa, que sustenta sistema de valores, relação hierarquizada de diferentes códigos em operação, segundo metassistema conceptual concernente ao sistema de significação produzido culturalmente e subjacente ao nível pré- e transcódigo da comunidade considerada.

Acrescentamos, ainda, que designações são produzidas, recortes são elaborados no *continuum* da existência (HJELMSLEV:1971), a partir de sistemas precedentes compatíveis e coerentes com a relação entre classes de equivalência e domínios de experiência, possíveis por uma metassemiótica cultural (estrutura estruturante) e construídos por um léxico e gramáticas conceptuais, que permitem a produção e o relacionamento de recortes culturais (RASTIER: 1987). A passagem de um código a outro se dá graças à intermediação desse metassistema conceptual, subjacente – em nível de estrutura hiper-profunda – em relação aos Discursos e Universos de Discursos e aos sistemas em constante operação. O metassistema conceptual apresenta-se, pois, como um primeiro nível de análise dos dados da experiência, configurado por um léxico conceptual, segundo regras de sintaxe conceptual de um lado, e um conjunto de enunciados conceptuais de outro, interdependentes.

Ora, isto conduz a pensar que a ligação entre *sémiosis* e *noésis* é muito mais estreita e mais profunda do que geralmente se admite. Representações semióticas e representações mentais não podem ser consideradas

independentes, embora pertençam a domínios diferentes. O desenvolvimento das representações mentais se efetua(m) como uma interiorização das representações semióticas ao mesmo título que as imagens mentais são uma interiorização de *perceptos* (VYGOTSKY:1985; PIAGET:1968^a; DENIS: 1980 apud DUVAL:1995). A compreensão conceptual, a diferenciação e a matriz de diferentes formas de raciocínio, assim como as interpretações de enunciados ligam-se, inteiramente, à mobilização e à articulação quase-imediata de muitos registros de representação semiótica, cuja representação depende de coordenação entre interiorização de *perceptos*, imagens mentais, interiorização de representações semióticas e de desenvolvimento de representações mentais efetuadas, elaboradas por um sujeito, imerso na pluralidade dos sistemas semióticos que permitem uma diversificação de representações de um mesmo objeto e o aumento das capacidades cognitivas dos sujeitos e, em consequência disso, de suas representações mentais, uma atividade cognitiva primordial.(BENVENISTE:1998)

Desta feita, na perspectiva da semântica interpretativa, e da semântica e mesmo da análise da inteligência artificial, algumas questões se apresentam, tais como:

- a) que é ler um texto?
- b) que instruções possibilitam atualizar o conteúdo de um texto, desde os traços de significado?
- c) como acessar o sentido?
- d) como representar o sentido ou os sentidos de um texto e descrever as relações de sentido?
- e) por quais operações de interpretação pode-se chegar a produzir seu conteúdo?

3. Lexicometria, representação gráfica do conhecimento e ensino de Língua(s)

A análise lexical, textual e discursiva, com o auxílio de instrumento da informática e instrumentos matemático e estatístico (estatística paramétrica), coloca-se à disposição de uma análise objetiva, descritiva e “científica”.

Do tratamento automatizado de um inventário lexicográfico, constituído de 27.950 palavras-ocorrências, atualizadas em discursos-concretos - as 180 redações de alunos de série final de ciclo fundamental de estudos, recolhidas ao longo de um ano de letivo, distribuído em início, meio e fim de processo de ensino-aprendizagem, de seis (06) diferentes escolas de ensino fundamental, representantes de três diferentes regiões da Capital do Estado de São Paulo – foi-nos possível configurar do vocabulário geral avaliado, o vocabulário de base ou comum, o vocabulário preferencial e o vocabulário diferencial, um arsenal relativamente poderoso que permite a seus possuidores enfrentar a agressão dos objetos e das situações nominalizadas, que, por sua vez, dependem de um vocabulário muito mais amplo, inscrito numa “memória” que faz aparecer, a todo momento, eventuais sujeitos à procura de predicação.

Do ponto de vista do processo, a análise quali-quantitativa de textos, com o auxílio da ferramenta STABLEX – instrumento desenvolvido por André CAMLONG,(1991), da Universidade de Toulouse – Le Mirail – CRIC – permitiu-nos trabalhar com indivíduos lexicais, constituintes do texto que veicula o discurso, reter números de base para operar com medidas comparativas e descritivas, a fim de se caracterizar topológica e combinatoriamente os elementos lexicais do corpus dado e delimitado, de que já resultou *parte* das reflexões sobre aspectos relacionados ao ensino do vocabulário em

Língua Portuguesa (língua materna) em escolas de ensino fundamental, apresentadas, por nós, em dissertação de doutoramento, na FFLCH/USP—Departamento de Linguística.

Dessa etapa, buscamos, agora, o aprofundamento da avaliação, porque, acreditamos, ser possível verificar a possibilidade da produção de sistemas axiológicos, e as performances emissivas e as receptivas dos participantes de uma mesma comunidade sóciolinguísticocultural para poder situar a rede das comunicações sociais generalizadas e a rede das formas e modos de seu exercício, caracteres asseguradores do estatuto social do indivíduo, segundo estratos e estados definidos por exercícios de correlação e covariação entre elementos. Por certo uma tipologia de Universos Semânticos deve, ainda, ser construída, uma vez que um mesmo sujeito assume papéis sóciossemióticos diferentes, porque pertencente a grupos sociais diferentes e participante de diferentes e várias situações de comunicação.

Se, por um lado, a lexicometria, pelo encadeamento de medidas e operações, oferece a imagem do discurso – o peso de cada um dos casos da matriz; as linhas de força que nos permite descrever qualidades comuns, preferenciais e diferenciais do vocabulário efetivamente empregado; avaliar a tendência (reta de Henri) e compreender a integridade do discurso pela operação relação/valor (estrutura ou não de conjunto, presença de vocabulário rico ou pobre, e sua tendência ao enriquecimento ou ao empobrecimento); por outro, a representação gráfica do conhecimento pode se beneficiar dessa análise lexicométrica para deixar transparecer propriedades e marcas características, dadas por graus de variabilidade, tolerância e liberdade, e por linhas de determinação.

Este estudo, em seu segundo momento, teve então, como ponto de partida o levantamento geral dos dados já recenseados e tratados, numa primeira abordagem, para,

desse mesmo levantamento, aprofundar ainda mais o estudo. E, num momento ainda posterior, apoiado nos fundamentos de Semiótica, Lexicologia e Semântica Cognitiva, explorar, de forma gradual e mais aprofundada, a manifestação sintagmática aliada à forma de organização discursiva (tipos e gêneros textuais), compreendendo-se unidades frásticas e transfrásticas. Segue-se o estudo da organização específica das unidades constituídas, e, simultaneamente, da conotação particular, determinada e delimitada de tal tipo de organização textual, de que um ponto de convergência se instala: um entremeado na polissemia de todos os instantes-significações pertencentes a *n*-metadiscursos, dos quais *planos isotópicos* de ordem lingüística – interdependentes e intercomplementares – apresentam-se sob a forma de enunciados qualificados e funcionais, evocadores, certamente, de uma espécie de rede (esquemas de entendimento), componente(s) de uma taxonomia gramatical, definidora de um conjunto de interrelações e de interações, que é necessário examinar, para que, em seguida, seja possível a construção, então, de mapas conceituais ou esquemas de entendimento, segundo metarregras interpretativas atualizadas e postas em jogo(s) de linguagem, posto que um texto é produto do encontro entre uma língua, um homem e uma sociedade.

Dadas as características da sociedade contemporânea, em que novas formas de autonomia, de cooperação e de trabalho são assumidas, assuntos relacionados à **Ciência e Tecnologias da Linguagem** têm recebido especial atenção. Progressos na área tecnicocientífica têm caracterizado a contemporaneidade, no quadro das Ciências e suas aplicações, considerados diferentes eixos temáticos, relacionados a múltiplos assuntos, particularmente vastos, em suas diversas perspectivas.

No tocante às questões relacionadas à Educação, área caracterizada por relações de

autonomia e cooperação, compartilhamento de saberes em situação de interatividade social, como processo permanente de construção e busca de saber e de saber-fazer, a presença de alternativas para dar conta de dificuldades, vivenciadas tanto por educadores como por alunos, têm motivado a adoção de novas possibilidades pedagógicas, que, contextualizadas, procuram a valorização da atividade humana para a melhoria das condições de vida do homem.

E no tocante, sobretudo, à Pedagogia de Línguas, o desenvolvimento de competências e habilidades cognitivas e metacognitivas requer, tanto do professor quanto do aluno, um conjunto de (novos) valores, relações, interações e papéis para compreender mecanismos e procedimentos de processamento e reconhecimento de regulamento e controle do próprio conhecimento, sob diferentes dimensões e formas de interação. E ou se considerem características e princípios relacionados à ordem natural do mundo, refletidos na linguagem, ou se considerem natureza e dimensões constitutivas da faculdade de simbolização, ou se considere o conjunto dessas relações e princípios, por sua vez, relacionados a sistemas de ensino versus sistemas de aprendizagem, estão no centro das discussões o homem, o pensamento e o simbólico; a imagem e o discurso no processo das relações de identificação, o conhecimento e o reconhecimento: o fio de Ariadne.

Diante do exposto, julgamos interessante o levantamento de aspectos que possam contribuir para a reflexão sobre a Pedagogia de Línguas, notadamente para o Ensino de Língua Portuguesa, considerados também os documentos oficiais, levando-se em conta a formação linguístico-cultural dos participantes de uma comunidade considerada e critérios lexicomáticos e conceituais, dentre outros, que possam orientar a leitura, a interpretação e a produção de textos, uma das áreas possíveis de aplicação. A partir de situação de ensino-

aprendizagem de língua já amplamente conhecida, constituem-se **objetivos**:

- a) mostrar a importância e o alcance de um estudo de texto, do ponto de vista da análise lexicográfica, como instrumento capaz de medir e mapear conhecimento para atingir propósitos produtivos;
- b) analisar os impactos dessa base tecnológica sob o número de aspectos de desempenho e formação de redes de inovação, indicativas de novas abordagens e procedimentos metodológicos.

E no conjunto dessas reflexões gerais, formalizam-se também:

- a) desenvolver índices adequados para a produção, seleção, armazenagem e atualização de conhecimento por meio de representações lexicais, textuais e discursivas;
- b) trabalhar com representações gráficas do conhecimento, no domínio da psicologia e da semântica cognitiva;
- c) revisar, de forma gradual e aprofundada, elementos de trabalho com vistas à descrição e atualização de instrumentos, mecanismos e procedimentos;
- d) avançar o estudo descritivo pela análise dos dados por meio de exercícios de correlação e covariação, através de procedimento da estatística paramétrica, objetiva, científica e indutiva;
- e) analisar o funcionamento de aplicativos disponíveis, e acessíveis, como ferramentas, possíveis, embora não imprescindíveis, de mapear conhecimento;
- f) padronizar documentação por meio de processo de automação, de criação, catalogação e reorientação com base em regras pré-estabelecidas;
- g) estruturar o conhecimento para representar um processo, a fim de se criar uma base de conhecimento;

h) categorizar e construir mapa(s) conceitual ou mapa(s) de competência;

i) promover considerações teóricas e práticas que possam advir como constructores e indicadores de inovação, com base em conhecimento e tecnologia apropriados para representar propriedades do conhecimento;

4. Esquemas de entendimento e representação gráfica do conhecimento

Representações mentais e representações semióticas não podem, pois, ser consideradas independentes. Não há *noésis* sem *sémiosis* - sem recorrer à pluralidade ao menos potencial de sistemas semióticos, recurso que implica sua coordenação pelo sujeito ele-mesmo. Isto nos conduz a pensar que há uma ligação estreita entre *noésis* e *sémiosis* e é sabido que a compreensão conceptual, a diferenciação e a matriz de diferentes formas de raciocínio, bem como as interpretações de enunciados ligam-se ao dinamismo e à articulação de muitos registros de representação semiótica. É, pois, uma atividade cognitiva primordial.

Trate-se de tarefa intelectual, de um exercício de raciocínio, de explicação ou cálculo, ou, ainda, de uma resolução de problema, ou de uma atividade de (re)textualização. Encontram-se, aí, representações semióticas convertidas para serem tratadas, cujo caminho consista em distinguir não só o valor epistemológico e valor de verdade de uma proposição, como seu estatuto e conteúdo, caminhando-se da análise funcional para a análise estrutural.

Nesta perspectiva, língua e fala apresentam-se como duas instâncias autônomas e instauradoras do lugar de encontro para a produção de sentido, que, aplicadas a uma comunidade cultural definida, colocam em evidência a distância separadora do que é coletivo e individual da existência semiótica. A resultante da descrição de

mecanismos da língua natural e da construção de conceitos epistemológicos que ela supõe confronta-se com todas as outras formas de existência social e individual.

Como já nos referimos, a compreensão conceitual, a diferenciação e a matriz de diferentes formas de raciocínio, assim como as interpretações de enunciados ligam-se, inteiramente, à mobilização e à articulação quase-imediata de muitos registros de representação simbólica, cuja representação depende da coordenação entre interiorização de *perceptos*, imagens mentais, interiorização de representações semióticas e desenvolvimento de representações mentais efetuadas e elaboradas por um sujeito – individual e coletivo – imerso na multiplicidade de sistemas de significação, que permitem a diversidade de formas de representação de um mesmo objeto e aumentam as capacidades cognitivas dos sujeitos e, em consequência, as suas representações mentais.

A forte ligação entre funcionamento cognitivo do pensamento e a diversidade de registros de representações, consideradas as duas atividades – o raciocínio e a compreensão de textos – poderia melhor descrever os processos que lhe são subjacentes? E será, pois, a forma figurativa – por meio da qual se manifesta a participação mais geral dos indivíduos que a compõem – uma das características principais da dimensão semiótica de uma sociedade considerada?

Considerando-se que as instruções interpretativas são construídas, e não dadas, a questão, então, poderia residir não na escolha deste ou daquele método, mas na estratégia adotada e na teoria de que ela procede. Desta feita, a definição da unidade lexical numa perspectiva enciclopédica, segundo uma norma regente, e a construção de isotopias, poderia permitir não só integrar aspectos importantes relacionados à extensão do corpus, ao conhecimento do sistema linguístico e estruturas idioletais e o sistema

funcional da língua, como também atualizar componente(s) semântico(s) do(s) texto(s) considerado(s), observando-se as diferenças diatópicas, diastráticas e diafásicas de uma língua funcional e sintópicas, sintrásticas e sinfásias, se considerado um só nível de língua e sob uma só forma de um só estilo.

De acordo com NOVAK (1981), mapas conceituais podem ser construídos de forma *uni*, *bi* ou *multidimensional*, de acordo com: 1. encadeamento linear de conceitos e, portanto, esquematização linear de idéias; 2. categorização de conceitos e ligações, segundo associações das mais simples às mais específicas, representadas por categorização causal ou temporal; 3. ou sob o formato de ‘hipertextos’. Mapa conceitual constitui-se em um conjunto de conceitos interrelacionados, segundo uma estrutura hierárquica proposicional, e permite, através de recursos gráficos, enfatizar semelhanças entre conceitos pela semiotização do percurso que conduziu à sua construção.

E segundo AUSUBEL (1963), mapas conceituais podem ser utilizados para a promoção de diferenciação conceitual, progressiva e integrativa, de sorte que cada indivíduo reelabora o processo de significação, atualizando sentidos em conformidade com sua própria bagagem cultural. Mapas conceituais aparecem também como uma tecnologia para negociar significados, o que é feito por meio de propriedades, tal que a forma de representação se fundamenta em:

- a) tipificação dos conceitos e classes de conceitos;
- b) níveis e operações de abstração envolvidos na categorização de informações.

Colocamos, mais ainda, a construção de uma microsemântica, como hipótese de trabalho, segundo a qual a construção de mapas conceituais, definidos como representações mentais, seriam organizados, de sorte que três percursos elementares de

atualização poderiam se realizados, a saber: 1. de traços de significado inerente a um outro traço de significado inerente; 2. de traço de significado inerente a aferente; 3. de traço aferente a um aferente, de acordo com norma(s) e lei(s) de interpretação contextuais à propósito da determinação da coesão textual.

Ora, se as condições de atualização de um traço de significado depende de uma pertença contextual, qual é a garantia desse pertencimento? Uma análise textual deve operar uma seleção por meio de traços atualizados, de que a estratégia e a tática interpretativas podem ser regradas por uma racionalidade, e ser o objeto de uma descrição científica, que, por sua vez, permita a plausibilidade de diversos percursos interpretativos, de sorte que se reconheça para cada palavra os componentes próprios de uma língua, que o contexto e a situação de comunicação convocam ou virtualizam. (RASTIER:1987b)

A construção de esquemas de entendimento e/ou a representação gráfica do conhecimento pode se beneficiar de estratégias e de procedimentos tais que os sentidos de um texto não sejam mais somente acessíveis à intuição, mas também, e sobretudo, descritos racionalmente.

5. Língua: sistema de possibilidades ou de virtualidades

A historiografia da Lingüística é marcada pela passagem de uma modelização referencial à intensional. A gramática da Antiguidade (perpetuada pelas gramáticas escolares) desenvolveu formas descritivas, de modo que categorias foram definidas em termos referenciais, sob tipologia diversa; categorias definidas, a exemplo, por referência a estados psicológicos dos locutores (modos optativo e dubitativo); por referência a tipos de objetos designados (nomes e pronomes definidos e indefinidos); por referência à posição da

própria categoria ou do seu substituto (advérbio, interjeição e pronomes substantivos); por relação ao contexto gramatical (particípio - nome/adjetivo e verbo - e conjunção - elementos de ligação de frases ou sintagmas) e por referência ao uso e à origem das formas (nomes - comum, próprio, de ornamento, e neologia - compostos e derivados).

A representação espacial permitiu a descrição de objetos e sua correspondência a dados exteriores; descrição, segundo uma justaposição, que definiu a língua como organismo estático, o que implica contexto referencial imediato e apoio em regras formuladas em termos de conteúdo, proposta de constituição hierárquica sobre regras, marcas flexionais e derivacionais, cuja base se dá sobre conteúdos nocionais e não de acordo com uma integração tipológica.

Disto dois fatos se apresentam estreitamente relacionados: o desenvolvimento de uma reflexão metodológica de um lado; e de outro, o desenvolvimento de um trabalho linguístico, que se portando sobre um objeto disperso, a história de uma língua ou a comparação de estados de língua, busca estudar marcas flexionais para a constituição da estrutura da palavra, pela busca de seu grau de composição, ou pela semântica das relação de actância ou por figuração morfogénica, sustentada pela polaridade emissão-recepção.

No entanto, um estudo científico da utilização da linguagem tem, como se sabe, seu ponto de partida num ato metafórico, que é um trabalho de conceptualização, em que são apresentados como objeto direto os enunciados teóricos, e como objeto indireto as suas aplicações, segundo setores específicos da realidade. Como exemplo, tem-se a criação de termos ou palavras técnicas, dada por modelo conceptual global, em que, por exemplo, um termo é criado por associação a um campo já conhecido, ou por modelo conceptual ativo em que o termo criado se dá segundo domínio adjacente, por meio da

associação de campos, ainda em termos espaciais, como, por exemplo, os termos ‘extensão ou ampliação de sentidos’ e ‘restrição ou deslocamento de sentidos’.

Desse aspecto, interroga-se sobre as expressões que, sem alguma relação significando substância do ser, são categorizadas em substância; quantidade; qualidade; relação; lugar; tempo; posição; posse; fazer, cuja organização retórica tem base semântico-gramatical, e constituem um quadro de descrição ou de exposição para a definição funcional, tal qual as gramáticas escolares. E, além deste, um outro fato é a disposição de metaforização interna (ou transferência interior) a partir da visada de uma língua, em que a explicação de universais linguísticos dá-se como correspondentes a um jogo de estratégias.

Do ponto de vista, então, de uma arqueologia do saber linguístico (Coseriu), um modelo teórico se insere num contato social, global e específico, cuja teoria deve ser definida em razão de outras teorias e modelos e cuja comercialização apresenta três componentes internos, a saber: a) formalização (ou conjunto de técnicas de descrição, propondo-se um inventário não só das regras descritivas ou eventualmente normativas, bem como dos constituintes, definidos como metarregras, em que regras codificam estruturas linguísticas; e constituintes, a representação formal de extensibilidade do modelo, cuja variação linguística deve ser explicada pelo próprio universalismo do modelo; b) conceptualização (ou nível de semântica articulado do modelo - nível de formação das noções teóricas ou estudo das macroestruturas do objeto descrito e articulação das unidades da descrição, bem como do processo típico de funcionamento da língua ou do discurso); c) auto-imagem (ou modelização do objeto e procedimentos da descrição e da relação entre objeto/sujeito; na linguística, o desdobramento do sujeito linguístico e sujeito da descrição).

Todavia, o campo metodológico coloca em foco questões epistemológicas, pois importam as condições de possibilidade, uma vez que no espaço do saber aparecem configurações que dão lugar a formas diversas do conhecimento empírico.

A linguagem, por sua vez, é determinada, na sua essência e manifestação, por dimensões universais, a saber: a) criatividade: atividade livre e criativa; b) materialidade: realização por formas materiais (fônica, gráficas etc); c) semânticidade: a toda forma corresponde um conteúdo e vice-versa; d) alteridade: manifestação primária do ser-com-outro, manifestação por outro e ao outro; e) historicidade: a linguagem manifesta-se sempre e necessariamente e sob a forma de uma língua, na tradição linguística de uma comunidade histórica.

A criatividade e materialidade são universais de toda forma de cultura, atividades criativas que se realizam por meio também de formas criativas; a semânticidade é diferença específica da linguagem dadas as formas sócio-culturais; e a alteridade é o traço distintivo do ato de significar linguístico em relação às outras formas expressivas e, conseqüentemente, o fundamento da historicidade da linguagem, manifestação constante da solidariedade com uma comunidade de sujeitos-falantes, com tradição e, pois, essência social e histórica intrínseca ao universo humano.

Desta feita, a linguagem - atividade histórica universal e realizada segundo as tradições das comunidades históricas - apresenta três níveis relativamente autônomos, a saber: a) nível universal do falar em geral; b) nível histórico das línguas; c) nível individual e particular do(s) discurso(s) ou discurso(s)-ocorrência(s) definido(s) como texto(s), unidade(s) de significação.

A estes três níveis correspondem três tipos de saber linguístico: a) saber elocucional ou saber em geral; b) saber idiomático ou saber falar tal ou qual língua; c) saber expressivo ou

saber construir tal ou qual discurso/textos. E a esses saberes correspondem também três tipos específicos de conteúdo: a) designação; b) significado; c) sentido. Aos quais, por sua vez, correspondem três tipos de conformidade do falar com o saber linguístico ao que se encontra correspondência a cada nível: a) congruência; b) correção; c) adequação.

Disto decorre que a oposição corrente entre linguística descritiva e linguística histórica é uma convenção terminológica, que se é possível até admitir não sem uma oposição entre descrição sistemática (ou estrutural) e história da língua. Entende-se o saber linguístico como um saber técnico, e que, nesse sentido, o saber idiomático são as técnicas históricas do falar. No entanto, correspondente a uma língua histórica e não a um sistema único e unitário; um conjunto de tradições em parte análogas e em parte divergentes.

Isto posto, a criatividade manifesta-se como variedade (considerando-se um estado de língua) e a alteridade como homogeneidade e unidade. Quanto à variedade linguística, três são os tipos fundamentais: a) variedade diatópica ou espacial; b) variedade diastrática ou sócio-cultural; c) variedade diafásica ou estilística. E são também, no sentido contrário, três os tipos de unidades, a saber: a) sintópica ou dialeto; b) sintrática ou níveis de linguagem; c) sinfásica ou estilos de linguagem.

O que é comum? É a língua funcional, um sistema delimitado cada vez mais como sistema homogêneo. Define-se uma língua histórica, então, como um conjunto complexo de dialetos, níveis e estilos (dado que o unitário só é segundo um só ponto de vista), ou um conjunto ainda mais complexo de línguas funcionais historicamente realizadas entre si.

Significa isto que, na prática, todo falante é um sujeito plurilíngue, porque conhece ativa ou passivamente, sob diferentes graus, mais de um dialeto e mais de um nível de língua e, por

definição, diversos estilos e, então, uma série de línguas funcionais, que efetivamente usa em seu falar, considerado um mesmo e único sistema linguístico que o sujeito-falante emprega em suas relações social e historicamente construídas.

Considera-se que um estado de língua é estático somente como uma técnica já realizada e na projeção sincrônica sobre um eixo imóvel, pois um sistema linguístico, de fato, se acha num duplo movimento: no sentido da variedade (difusão e recessão) e no sentido da renovação dos sistemas idealmente unitários. (COSERIU:1992)

Desse ponto de vista, uma língua não é algo já feito, acabado (*ergon*), mas, ao contrário, essencial e fundamentalmente, um sistema de modos de agir: uma prática ou técnica aberta, e, em consequência, um sistema de possibilidades ou de virtualidades que se realizam progressivamente no eixo da história; um saber-fazer produtivo, que apresenta, por sua vez, três níveis: a) sistema materialmente realizado e tornado tradição tal qual uma norma; b) conjunto dos modos de fazer funcional, tal qual um sistema; c) conjunto coerente das categorias funcionais e dos tipos de procedimentos materiais, tal qual tipo linguístico.

Assim, uma língua não é jamais feita: cria-se, faz-se e refaz-se continuamente (ainda que sem mudança aparente). Criação e recriação das tradições históricas linguísticas e emergência da língua como uma tradição, de modo que o ser da língua coincide com o vir-a-ser histórico. A redução convencional de uma língua a um só sistema linguístico - língua funcional- entende-se não só como uma sincronia ou como uma descrição sistemática ou estrutural e como uma diacronia ou estudo do desenvolvimento histórico de uma sincronia (ou gramática histórica). mas também como história da língua.

Nesse último sentido, os termos não são correlativos, posto que à sincronia corresponde não somente a descrição de um

sistema unitário mas também a descrição dialetológica, sociolinguística e estilística da língua. À descrição, então, corresponde um sistema não só sincrônico, mas sintópico, sintrástico e sinfásico.

Estabelece-se que a sincronia está para funcionamento, assim como mudança está para diacronia, um só momento, uma vez que as mudanças são “internas”, resultantes da dinamicidade própria da técnica idiomática, das realizações de virtualidades já dadas como tais: o que é mudança da norma é funcionamento do sistema e o que é mudança do sistema é funcionamento do tipo linguístico.

A norma se modifica exclusivamente pela aplicação do sistema (real ou re-interpretado); o sistema, pela aplicação do tipo ou pela aplicação do sistema mesmo como uma reinterpretação; e o tipo, somente pela reinterpretação, motivada ou não pela co-presença de dois ou mais tipos na mesma língua histórica ou por influência de sistemas linguísticos diferentes.

Diante do exposto, só a história pode corresponder o ser real da língua, considerando-se:

a) variedade e homogeneidade da língua - a descrição sistemática não pode considerar a variedade sem perder a sua coerência, já que não é bem secundário, mas um bem intrínseco à língua, refletor da criatividade essencial da língua, dado que uma língua não funciona somente por meio de estruturas internas, mas também pelo conhecimento dos falantes;

b) tensão entre estruturas diacrônicas correntes (virtuais e nova) , já que descrição sistemática de estruturas aparece em sistemas diferenças, ou variantes não-funcionais (ou estruturas opositivas diferentes);

c) processo de permanente constituição: realização progressiva na/da norma e do/no sistema de um sistema linguístico já

dado tanto quanto de sistema como de tipo.

A descrição sistemática considera a língua como uma técnica aberta ou um sistema de possibilidades, e, portanto, tenta dar conta da dinâmica interna própria de toda técnica idiomática. A história pode registrar a realização efetiva das formas e construção supostas como possíveis no momento da descrição e confirmar também sua realidade.

Considera-se a língua em suas relações com outras formas de cultura espiritual e material da comunidade falante e pode também dar conta de tudo isso que no falar (ou como conteúdo ou como estrutura sintática) é determinado não como um saber idiomático mas como um saber extralinguístico. Uma descrição sistemática limitar-se-ia à língua em si mesma, e, assim, ou interpreta a língua numa perspectiva deformada ou renuncia as determinações extralinguísticas do falar.

Na realidade, uma descrição - estrutural e funcional - não se acha fora da história, pois toda descrição de um objeto é parte da história, pois que dela não se sai. A história é, pois, ciência linguística integral, que aspira considerar seus objetos, por exemplo, a língua, sob todos os seus aspectos e com todas as suas determinações históricas internas e externas: ciência das línguas e história linguística.

Descrever, então, uma estrutura é dar conta das dependências e independências, sem ignorar a problemática semântica, porque não se pode reduzi-la a apenas um repertório de significações diferentes, posto que uma variante é sempre função de uma forma; uma significação particular é sempre função da significação fundamental. deduzida logicamente de significações particulares, posto que a própria função constitui o princípio da dependência, inerente e constitutivo da estrutura. (HJELMSLEV:1971)

6. A análise lexical, textual e discursiva

A análise lexicométrica, auxiliada por ferramenta tecnológica, no caso deste estudo pela ferramenta *STABLEX* (CAMLONG & BERTRAND - *Stablex* - Software PC2004), permite descobrir co-relacionamentos e dados implícitos nos registros de um corpus, pelo estudo e desenvolvimento de um processo de extração, pela:

- a) detecção de dependências entre os dados, para se identificar os atributos e suas relações de interdependência;
- b) detecção de desvios para que se identifiquem elementos que se encontram fora dos padrões estabelecidos ou esperados;
- c) identificação e análise de estruturas e de objetos com características comuns;
- d) descoberta de fórmulas ou modelos capazes de descrever os objetos ou conceitos envolvidos.

Para tal, constrói-se uma metodologia adequada, que compreende etapas de pré-processamento (mineração que inclui processos de análise, integração, re-informações e limpeza de dados); processamento (processamento de dados, predição, regressão, classificação, constituição de conglomerados, associações, visualizações), e pós-processamento (aplicação de filtro, seleção e ordenação das descobertas, elaboração de mapeamentos de representação e elaboração de relatórios para a apresentação da base de conhecimento).

A base de dados constituiu-se de redações de alunos da última série do ciclo fundamental de ensino, de seis diferentes escolas da rede estadual paulista, selecionados aleatoriamente dez alunos de cada uma das seis escolas, cujos textos foram produzidos em três momentos diferentes do processo de ensino-aprendizagem, totalizando, então, para cada escola, três diferentes momentos da produção

textual dos alunos - período inicial, medial e final do processo de ensino-aprendizagem de um ano letivo.

O estudo não aponta considerações especificamente sobre as escolas, mas tendências, advindas da análise evolutiva da produção textual escrita dos alunos, que podem, por sua vez, representar aspectos a serem considerados em uma discussão sobre ensino de línguas, sobretudo, ensino de língua materna.

Os textos coletados, segundo os três diferentes períodos do processo de ensino-aprendizagem durante um ano letivo (períodos inicial, medial e final coincidentes a início, meio e final do ano letivo, respectivamente), receberam tratamento especial para a classificação dos objetos (palavras), desde a aplicação de técnicas de análise de informação de dados lexemáticos para a identificação de padrões e relacionamento entre os objetos, de sorte que os textos, depois de tratados, puderam ser convertidos em formato TXT ou ASC II, que é semi-estruturado ou desestruturado por apresentar características complexas.

Deve-se estabelecer, aqui, uma diferenciação entre classificação de dados e a descoberta de aglomerados (grupo de entidades similares ou regiões contínuas de um espaço contendo uma alta densidade relativa de pontos, separada por outras regiões como esta por regiões contendo baixa densidade relativa de pontos).

A classificação de dados ocorre por conhecimento do próprio objeto, pela possibilidade de se estabelecer descrição para a classe dos objetos, segundo uma análise do objeto e atribuição do objeto a uma classe, o que é, pois, uma definição apriorística.

Diferentemente desta metodologia, operou-se com a descoberta de padrões, e de associações entre os padrões para, daí, efetuar-se a definição de classes ou de categorias, uma vez que não se conhecia de antemão a classificação a que os objetos (dados) seriam

rotulados. Identificou-se a melhor distribuição ou configuração, não pela comparação dos dados a objetos já conhecidos, mas com base na identificação de similaridade entre objetos segundo atributos em comum, cujas variáveis analisadas geraram grupos homogêneos segundo análise combinatória e de alta complexidade.

Do corpus tratado, as primeiras extrações ou a ordenação dos dados foram obtidas: a confecção do léxico (léxico alfa e léxico delta) que consistiu na disposição das palavras para a organização das ocorrências segundo o critério da frequência, a primeira etapa do trabalho, e da indexação das palavras, segundo a frequência absoluta, a que se seguiu a apresentação da tabela de frequência, e desta as operações subsequentes (tabela de desvios reduzidos e tabela dos valores lexicais) .

6.1. Representação de objetos ou padrões

O esquema representativo levou em conta a representação por vetores de palavras (seleção e extração de características) pelo critério da **seleção de características**, segundo representabilidade, efetividade e relevância, em que a frequência absoluta (números de vezes em que a palavra ocorre num corpus) é uma medida importante, mas medida insuficiente. É ainda necessário estabelecer-se a frequência relativa dada pela relação de uma palavra a todas as outras do documento considerado, no caso deste estudo, o agrupamento das redações por alunos/série/ período/escola/região. Necessita-se identificar o peso ou a força de representabilidade de cada palavra, para se operar uma normalização do número de ocorrências em relação ao número total de ocorrências de todas as palavras presentes nos documentos considerados, justamente porque palavras não ocorrem eventualmente em um texto.

A repetição de uma palavra por parte do enunciador pode ser um indicador da significância das palavras-ocorrências, pois a

inclusão de um elemento (novo) em enunciados demanda um esforço maior, e, por isso o enunciador lança mão da repetição de palavras, enquanto avança ou varia sua produção textual. Tais medidas de frequência - absoluta, relativa e inversa - selecionam as palavras discriminantes em algoritmos de classificação ou de recuperação de informação com perspectivas de identificação de centróides ou matrizes; o corte se faz a partir do próprio peso, pela estipulação de um valor mínimo e de um valor máximo.

Esta etapa do estudo trabalhou com a confecção de tabelas de frequência e de desvios reduzidos. A Tabela de Desvios Reduzidos é considerada etapa fundamental para os processos posteriores, que consistem nos cálculos algébricos, a análise paramétrica propriamente dita. A extração de dados ou de informações é uma técnica importante que, ao utilizar algoritmos e ferramentas da área, objetiva a extração de dados e informações contidas nos textos para dar conta da extração de identidade de dados e informações relevantes.

No caso deste estudo, a relevância das imagens apontam informações sobre a evolução da aprendizagem no processo de ensino, considerados os conglomerados dos textos elaborados por alunos das diferentes seis escolas em seus respectivos períodos de aprendizagem (1- início; 2- meio e 3- período final de um ano letivo), como demonstram os **gráficos 1 a 6** correspondentes ao peso lexical apresentado pelo conjunto de texto produzidos pelo grupo de alunos/série final de ciclo fundamental de cada uma das seis escolas, fruto da análise paramétrica:

Gráfico 1 – escola 1

Gráfico 2 – escola 2

Gráfico 3 – escola 3

Gráfico 4 – escola 4

Gráfico 5 – escola 5

Gráfico 6 – escola 6

A análise paramétrica, importante ferramenta de análise e síntese – objetiva, indutiva e científica - possibilitou observar, decompor e manipular dados, com a ajuda do software *STABLEX* (CAMLONG & BERTRAND:2004), e chegar, pela arquitetura dos textos, à composição do discurso. A análise da representação da situação textual deu-se de forma controlada pela própria especificidade, por planos de produção e compatibilizações sêmicas e semêmicas (BARBOSA:1996), em que elementos linguísticos foram reduzidos a dados mensurados e proporcionalmente dimensionados (estudos de frequência e desvios reduzidos), de tal modo que foi

possível examinar o índice nocional, quantitativo e qualitativo.

Para a produção de novas características mais salientes, ou mais discriminantes ou mais descritivas, criaram-se novos descritores, quer pelo mapeamento de conceitos, em que palavras são transformadas em estruturas mais complexas, quer pela análise fatorial em que as palavras são reduzidas à dimensionalidade, agrupando-se variáveis correlacionadas.

Essa operação conta com a medição da proximidade, associação ou semelhança entre padrões, que possibilitou a descrição dos documentos por valores qualitativos e por valores quantitativos. Tais valores de proximidade (proximidade, associação e semelhança) correlacionam-se à distância (física, considerado o espaço euclidiano) entre os objetos, em que, na operação, por se relacionar mais ao valor quantitativo, o conceito de distribuição foi substituído pelo conceito de similaridade, dado por medidas de distância, coeficientes de correlação, coeficientes de associação e medidas probabilísticas de similaridade.

Foram valores estabelecidos segundo uma **métrica**, medida estabelecida de forma objetiva, reprodutível e quantificada, de acordo com a natureza das características do objeto (palavra): simetria, densidade triangular, densidade de não-identidades e indistinguibilidade de identidades (ou equivalentes). A medida de distância é importante para a verificação da distribuição dos objetos em espaço euclidiano, em que cada ordenada corresponde a uma característica. A similaridade é uma operação dada pela proximidade, associação e semelhança das palavras (objetos) nesse espaço.

A etapa seguinte compreendeu a análise do agrupamento de entidades similares, ou análise de regiões contíguas de um espaço, contendo alta densidade relativa de pontos, e separada por outras regiões contendo baixa densidade relativa de pontos. Operação que se

caracteriza por densidade, por variância (o grau de dispersão dos pontos presentes no agrupamento, em relação ao seu centro, ou os pontos (ocorrências) em relação umas às outras), por dimensão (indicativo do raio ou diâmetro, indicativo do conceito de conectividade entre os pontos de um conglomerado) e por separação (grau de sobreposição ou isolamento dos conglomerados no espaço),

6.2. Identificação e análise dos agrupamentos utilizando-se de conceitos

A adequação de uma metodologia é fundamental para a identificação e análise de agrupamentos, decorrendo de que apenas o critério da frequência é insuficiente, porque apenas dá conta de palavras soltas e ainda descontextualizadas, de sorte que nem idéias, nem objetos e nem ações são asseguradamente identificados, por não se oferecer a seleção de representações e discriminações adequadas.

Para minimizar a questão, a lexicometria oferece técnicas, como as sumariamente descritas, que possibilitam um estudo controlado e padronizado, de que a partir do conjunto de textos de alunos (conglomerados) foi possível para cada um dos agrupamentos estudar os valores lexicais, e obter o Vocabulário Preferencial, Vocabulário Comum (de base positiva e de base negativa) e Vocabulário Diferencial, caracterizadores e identificadores de aspectos da produção textual dos alunos não só em cada um dos três períodos, como na globalidade do processo evolutivo do processo de ensino-aprendizagem, considerados cada um dos grupos de alunos de cada uma das seis escolas durante um ano letivo.

Para, apenas, uma amostragem, os gráficos 7, 8 e 9 apresentam os dados - relacionados tão-somente e apenas à Escola 4, de cujos valores o leitor poderá fazer sua própria análise:

Gráfico 7
(período inicial)

Gráfico 8
(período medial)

Gráfico 9
(período final)

escola 4 - períodos e valores lexicais correspondentes, considerando-se início, meio e fim do processo de ensino-aprendizagem durante um ano letivo:

Há que se considerar, para a análise dos agrupamentos, a presença também de descritores específicos, como, por exemplo, a presença de *hápax* ou ocorrências altamente discriminantes de um dado aglomerado, que se caracterizam pela especificidade, pois se apresentam uma única vez no conglomerado e

especificamente num determinado e único aglomerado, sendo, pois um forte índice de significância de aspectos identitários do próprio aglomerado em que ocorrem.

Há também outras ocorrências caracterizadoras de aglomerados, como as ocorrências que aparecem x^{+1} no conglomerado e apenas num aglomerado específico, mas, aí, ocorrendo, no entanto, apenas uma única vez.

7. Conclusão

Vive-se hoje num ambiente sistêmico, complexo e dinâmico. Há um volume de informações que a recuperação ou a identificação de informações relevantes e não só a análise e a manipulação bem como a descoberta de novas informações não se constituem tarefa das mais simples ou fáceis. Nesse sentido o processo da descoberta de conhecimento apoiado por ferramenta tecnológica é um processo de análise de dados ou informações para a aquisição de novos conhecimentos a partir da manipulação de uma grande massa de dados ou informações.

Da análise de frequência ao cálculo algébrico, foi-nos possível a observação e o estudo de tendências e convergências que enriqueceram a compreensão tanto da natureza dos elementos codificados quanto do cruzamento das variáveis contextuais. Da análise lexical e textual determinaram-se valores, hábitos e força de certas cargas emotivo-afetivas, e evidenciaram-se operações lógicas da superfície discursiva e redes

produtoras de efeitos de sentidos. A quantificação do vocabulário tornou possível atingir três níveis fundamentais da linguagem: **vocabulário maciço ou preferencial** ($z \geq +1,96$); **vocabulário comum, base média** ($-1,96 \leq z \leq +1,96$; e **vocabulário único ou diferencial** ($z \leq -1,96$), para cada uma das seis escolas em seus respectivos períodos de ensino-aprendizagem. (z = desvio reduzido)

Determinaram-se valores, hábitos e força de certas cargas emotivo-afetivas, evidenciaram-se operações lógicas da superfície discursiva e redes produtoras de efeitos de sentidos, atingindo-se três níveis fundamentais: **vocabulário maciço ou comum** ou zonas de intensidade, conformismo, clichês e estereótipos lingüístico-culturais; **vocabulário comum** ou base média sobre a qual se estabelecem a comunicação e o universo mental, implicações de um dado grupo social e **vocabulário único ou diferencial**, carregado de informações, esperas e desventuras.

A identificação de redes de conhecimento, dimensão de análise transversal torna possível a revelação de perfis importantes para redefinição de pré-requisitos e exigências para desenvolvimento de competências e, conseqüentemente, esforços num trabalho conjunto, crucial no processo de interação, inovação, desenvolvimento e gerenciamento do potencial humano, em que novas formas e abordagens de socialização do conhecimento são indispensáveis às novas arquiteturas de aprendizagem.

Referências Bibliográficas

- AUSUBEL, D.. P. **The psychology of meaningful verbal learning**. New York, Grune & Stratton, 1963.
- BARBOSA, M. A. **Língua e discurso: contribuições aos estudos semântico - sintáticos**. SP: Ed. Plêiade, 1996.
- BENVENISTE- E. **Problemas de Linguística Geral I e II**. UNICAMP:Ed. Pontes:1998.

- CAMLONG, A . & BERTRANS, T. **STABLEX - Version PC- Université de Toulouse II** - The Art Editora Ltda, Pirus Tecnologia., 2004.
- CAMLONG, A . **Méthode d'analyse lexicale, textuelle et discursive**. Paris: OPHRYS, 1996.
- _____. *Lexicométrie I, II, III, IV, V e VI* - In.: **STABLEX – Version PC- Université de Toulouse II**- The Art Editora Ltda, Pirus Tecnologia, 2004.
- COSERIU, E. **Communication & Cognition**. V.25, n° 213, 1992.
- DUVAL, R. **Sémiosis et pensée humaine – registres sémiotiques et apprentissages intellectuels**. Bern, Berlin, Frankfurt/M. New York:1995.
- HJELMESLEV. **Essais Linguistiques**. Paris:Les Editions Minuit, 1971.
- NOVAK, J. D. **Uma teoria de educação**. S.P.:Editora Pioneira, 1981.
- PAIS, C. T. **Conditions sémantique-syntaxiques ET sémiotiques de La productivité systémique, lexical et discursive**. Thèse présentée à l'université de Paris IV – Doctorat d'Etat, avril, 1993.
- RASTIER, F. **Sémantique Interprétative**, Paris: P.U.F, 1987.
- RASTIER, F. Sémantique et intelligence artificielle. In.: **Revue Langages**, Larousse, sept/1987b.